Report No. FAA-RD-80-68

Aircraft Alerting Systems Standardization Study

D6-49273
February 1980

PHASE I

FINAL REPORT

Document is available to the U.S. public through
the National Technical Information Service,
Springfield, Virginia 22161.

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
Systems Research & Development Service
Washington, D.C. 20590

14 DG-49273

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
18 FAA-RD-80-68 ✓	AD-A090834	

4. Title and Subtitle	5. Report Date
6 Aircraft Alerting Systems Standardization Study. ~~Phase I Final Report~~	11 Feb[redacted]
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
10 G.P. Boucek J.B. Erickson B.L. Berson D.C. Hanson M.F. Leffler [redacted]	

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Systems Technology, Research - Crew Systems Boeing Commercial Airplane Company P.O. Box 3707, Seattle, Washington 98124 Douglas Aircraft Company - Lockheed Aircraft Company	
	11. Contract or Grant No. 15 DOT/FA79WA-4268

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Systems Research and Development Service Federal Aviation Administration Department of Transportation, Washington, D.C. 20590	9 Final Report [redacted] I Jan[redacted] - Novem[redacted]979 on
	14. Sponsoring Agency Code ARD 312 Phase 1

15. Supplementary Notes

17 80-68 12 372

16. Abstract

This report is one of a series of documented studies and experiments directed to the improvement and standardization of cockpit alerting systems. The efforts in this current study are noteworthy in that the three major manufacturers of Commercial Transport Aircraft, Douglas, Lockheed, and Boeing conducted the tests and co-authored this report. The primary purpose of the study is to extend and validate, through simulation, the precepts advanced in the previous contract study. The primary purpose of this report is to document the results of objective and subjective tests and the development of the functional design of the candidate alerting systems which will be implemented and evaluated in a flight simulator.

17. Key Words	18. Distribution Statement
Alert; Caution/Warning; Warning Systems; Stimulus Response; Human Factors; Alerting Systems; Speech Aircraft; Signal Response; Visual Alerts, Intelligibility; Signal Detection	Document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	370	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

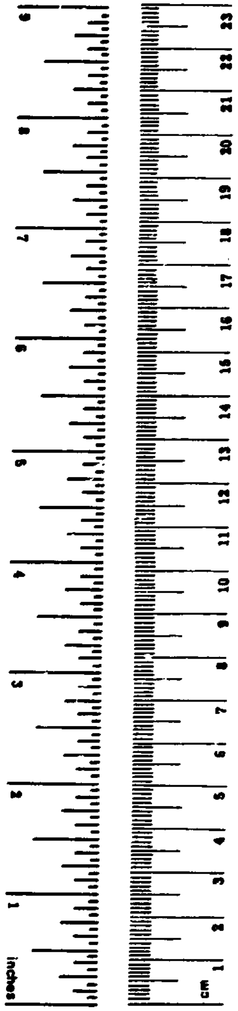
390145 Gm

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

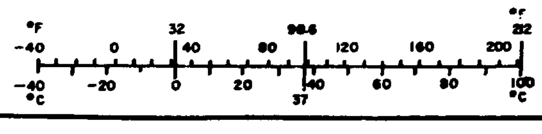
Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in = 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



PREFACE

This report is one of a series of documented studies and experiments directed to the improvement and standardization of cockpit alerting systems. The efforts in this current study are noteworthy in that the three major manufacturers of Commercial Transport Aircraft, Douglas, Lockheed, and Boeing conducted the tests and co-authored this report. The primary purpose of the study is to extend and validate, through simulation, the precepts advanced in the previous contract study. The primary purpose of this report is to document the results of objective and subjective tests and the development of the functional design of the candidate alerting systems which will be implemented and evaluated in a flight simulator.

The authors want to express appreciation to the many pilots of the three companies who participated in the tests. Also, the experience and guidance of Wayne Smith, the Boeing Program Manager, was of great value, as were the contributions of Dick Gabriel and Art Torosian of Douglas, and Ralph Cokeley, Les Susser and Chuck Mercer of Lockheed. Thanks go also to Russell White for his efforts in both preparing the simulator and conducting the tests. The contract sponsor is the FAA, and technical guidance was provided by Mr. John Hendrickson, ARD 312, the contract monitor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADI	Attitude Director Indicator
ALPA	Airline Pilots Association
ANOVA	Analysis of Variance
ARP	Aerospace Recommended Practice
ATC	Air Traffic Control
CAWS	Caution and Warning System
X^2	Chi Squared
CPU	Central Processor Unit
CRT	Cathode Ray Tube
dB	Decibel
DETAC	Digital Equipment Technology and Analysis Center
df	Degrees of Freedom
ELF	Electroluminescent Film
ft-L	Foot-Lambert
HUD	Head-Up-Display
Hz	Hertz
ILS	Instrument Landing System
LAX	Los Angeles International Airport
LCD	Liquid Crystal Display
LED	Light Emitting Diode
mL	Millilambert
msec	Millisecond
NASA	National Aeronautical and Space Administration
PAWS	Phase Adaptive Warning System
sin	Sine of an Angle
SMOLD	Switch Monitor and Light Driver
S/N ratio	Signal to Noise Ratio
VFR	Visual Flight Rules

1.0 INTRODUCTION

This contract is the fourth in a series having evolved from individual study efforts of independent altitude monitor requirements, and cockpit alerting system criteria, to this study which will develop standardized requirements for cockpit alerting systems. With each effort it became increasingly obvious that the design of aircraft alerting systems had not followed a systematic apprcach but rather had been detailed by the requirements of individual systems. While a good data base of philosophies, response characteristics, and guidelines was obtained, increasing evidence of alert proliferation and inconsistent application of alert concepts was found. The results of a contract issued to study the entire cockpit alerting problem and reported in FAA-RD-76-222, verified the finding of proliferation and inconsistencies. This study clearly indicated that the aircraft manufacturing industry needs a new set of guidelines for designing future cockpit alerting systems. As a result the study provided general recommendations for standardization of alerting functions and methods. However, since the recommendations were based on subjective conclusions drawn from an extensive literature search and detailed surveys of aircraft design and operating manuals, it was concluded that the recommendations required considerable refinement and ultimate evalution and validation.

The current contract designed to obtain this refinement and validation consists of three phases. The first phase evaluated the elements of alerting systems and provide candidate systems for evaluation; the second provided a detailed test plan for evaluating these candidate systems. The third phase will have line-qualified pilots exercise the candidate systems in a fixed-base simulator and will generate a set of alerting system design guidelines based on the results obtained from the evaluations. The guidelines will not define a single alerting system design that each manufacturer must utilize, but rather will provide criteria for numerous acceptable system concepts that will promote functional standardization. These guidelines will have been substantiated by experimental data and will reflect the concensus of the commercial transport aircraft manufacturers, certificating organizations, operators and pilot organizations.

The study is being conducted as a joint effort between the three major commercial airframe manufacturers, Boeing, Douglas, and Lockheed, and is sponsored by the FAA.

This report documents the conduct of phase one, which is the evaluation of the elements of alerting systems and the formulation of candidate system designs.

2.0 EXECUTIVE SUMMARY

At the beginning of this first phase of the study, after the contractual procedures were completed, the representatives from the three participating companies met to lay out program activities. The objectives and contractual requirements were outlined, and to meet the objectives, several supporting requirements were defined:

- Establish ground rules for the conduct of the study
- Establish assumptions for basic system design
- Determine what questions/issues needed to be resolved in Phase I
- Determine variables for Phase I experimental tests
- Conduct tests as necessary to obtain system component characterization
- Formulate candidate system concepts from viable system components
- Develop a test plan to validate system concepts
- Develop simulation capability for the selected system concepts
- Perform comparative simulator evaluations of selected system concepts
- Develop design guidelines for a standardized alerting system
- Assess the certification impact of the alerting system guidelines on aircraft design

After initial meetings which established the working group, overall objectives, ground rules, and a schedule of activities, candidate system concepts for Phase I were formulated. Five tests were designed to obtain needed data on several component variables. The independent variables were broadly differentiated as visual system elements and aural system elements.

The dependent variables (the performance measures) were defined to be detection time, response time, and response accuracy, response sequence, secondary task performance (visual, auditory), and subjective evaluations. The first two tests concerning visual system elements were conducted at Boeing; the second two tests involved aural system elements and were conducted at Douglas. Test five, a subjective test, was conducted at both Boeing and Douglas. Pilot-subjects came from all three companies, Boeing, Douglas, and Lockheed.

This section presents a detailed description of the major findings, and conclusions of the Phase I effort. The areas of work that will be covered in this summary will be the literature review, the test data and candidate system development.

2.1 LITERATURE REVIEW

A literature review was conducted to identify factors that influence the effectiveness of visual and auditory aircraft alerting signals. The major findings are summarized below and described in detail in section 3.0.

2.1.1 CONVENTIONAL VISUAL DISPLAYS

1. Location - High priority signals should be located within 15° of the pilot's centerline of vision. Lower priority signals should be located within 30° of the pilot's centerline of vision.

2. Size - High priority visual signals and alphanumeric legends must subtend no less than 1° of visual angle, lesser priority signals no less than 0.5°.

3. Brightness - High priority signals must be at least 10% brighter than surrounding signals and independent of aircraft ambient light conditions.

4. Steady state or flashing - The fastest mean detection times are obtained for flashing signal lights against a steady background; an ideal visual system would therefore flash the warning light with all background lights steady state, either on or off.

5. Color - Based on empirical data and cultural expectations the following color coding scheme should be used for visual aircraft warning and caution signals:

Red	- Highest priority signals - warnings
Amber/Yellow	- Lower priority signals - cautions
Green, blue, or white	- Advisory or informative data

2.1.2 ADVANCED DISPLAY TECHNOLOGIES

1. Overall evaluation - Since programmable display requirements are anticipated, Cathode Ray Tub (CRT) technology is recommended. Flat panel displays have progressed significantly during the past few years; however, factors such as luminous efficiency, resolution, multicolor capability brightness and cost do not yet favorably compare with the CRT.

2. Cathode ray tube - The CRT offers numerous advantages, such as: high resolution, good addressability, high contrast, flexibility, multi-color capability and relatively low cost. Its main disadvantages include: large volume, high voltage requirements, and limited useful life under high ambient light conditions.

3. Light Emitting Diode (LED) - The main advantages of LEDs are their extremely fast switching speed. LEDs are available in several colors (red, green, yellow), and are relatively inexpensive. A disadvantage is the higher power required, compared to other flat panel technologies.

4. Plasma Displays - The major advantages of plasma displays include: inherent memory (which eliminates the need for a fast refresh buffer and high bandwidth communication), small volume, no flicker or jitter, high resolution matrix addressing, high contrast ratio, and rugged construction. Disadvantages include relatively poor capability in high ambient light environments,

5. Electroluminescent Film (ELF) - The major advantages of ELF displays are ruggedness and large size potential, high contrast and uniformity of

brightness across the display surface. Its disadvantages include moderate brightness and moderate resolution capability.

6. Liquid Crystal Display (LCD) - Since passive LCDs reflect light rather than emit, they provide a significant advantage in high ambient light environments, for they allow a high contrast ratio to be maintained at all times. Other advantages include low switching voltage and high resolution. The major disadvantages include slow switching speed, line-at-a-time addressing, and cost.

2.1.3 AURAL ALERTING SIGNALS (NON VERBAL)

1. Frequency - It is important that no signalling device use a single frequency, but rather they should be a combination of sounds. To maximize perceived loudness a midfrequency tone (2000 to 4000Hz) should be used To enhance signal detectability sounds with frequencies of 250 to 4000 Hz should be used.

2. Intensity - The sound level of alerting signals should be 5 to 10 db above the masking threshold.

3. Location - The alerting signal source should be separated at least 90° from the source of interfering noise or messages. In addition, monaural signals should be presented to the dominant ear.

4. Message Content - Intermittent signals are more likely to be detected than steady-state signals, and should therefore be used for alerting signals.

5. Number of messages - The number of aural tones in the cockpit should be minimized. It has been recommended that no more than four distinct tones be used.

6. Environmental Factors - Since the detectability of alerting signals in the presence of distracting signals is enhanced by bimodal signal presentation, a combination of visual and auditory signals should be used for high priority alerts.

2.1.4 VOICE WARNING SIGNALS

Speech generating technique - Data indicates that voice modeling (prerecorded) presentations are more intelligible than phoneme-synthesized verbal messages. The voice modeling approach is more acceptable to pilots since it more closely resembles every day speech.

1. Alerting tone requirements - The question of whether a non-verbal tone should precede verbal messages remains controversial. Empirical data indicates that the use of a precursor tune increases pilot response times, and does not tend to reduce error rates (missed alerts). However, the majority of pilots surveyed favor precursor tones.

2. Voice model - The major factor in selecting a voice model (either male, female or electronic) is its intelligibility in the cockpit. The voice model selected for cautions and warnings should be qualitatively different from other voices in the cockpit. The results of several surveys indicated a strong preference for female voice models, while empirical data is ambiguous.

3. Voice inflection - Due to the potentially distracting effects of urgent-sounding verbal messages, a monotone inflection should be used for all voice messages.

4. Message format - Highly discriminative keywords or phrases should be used for all verbal messages. Whereas a sentence structure format has been shown to be more intelligible and to decrease response times, the majority of pilots surveyed prefer the keyword format. Also the advantages of the sentence format may disappear if a precursor tone or word is used.

5. Intensity and control - Verbal messages should be 5 to 10 db louder than the masking threshold, and the intensity should be adjusted automatically as the sound level changes in the cockpit.

6. Masking effects - The content of verbal messages should be designed to minimize potential masking effects produced by other sounds in the cockpit (aircraft noise, messages, etc.).

7. Mode of presentation - Verbal messages should be presented via speakers that are at least 90^0 from interfering speakers.

2.1.5 CREW OPTION AND CONTROL

1. Inhibition - Inhibition refers to delaying the onset of noncritical alerts until after critical or high workload flight phases. Pilot preference data indicates that non-critical alerts should be inhibited. However, little empirical or analytical data exists on what alerts should be inhibited or how the alerting system inhibit scheme should operate.

2. Prioritization - Prioritization of alerts consists of grouping alerts into criticality categories (e.g., warning, caution, advisory) and then evaluating the importance/urgency of these alerts within categories. Survey results indicate that the vast majority of pilots feel that alert effectiveness could be improved by prioritization. Analytical results, however, indicate that much more effort is required to develop useful prioritization schemes, particularly in the event of multiple failures, and that prioritization schemes should be developed for only the highest levels of alerts (warnings and cautions).

3. Store and recall - Store and recall refers to the clearing and subsequent redisplaying of uncorrected fault indications. Survey results indicate that pilots should be able to store/recall uncorrected fault messages. Most pilots indicated that warning messages should not be manually cancellable, others stated that they should, but that more stringent criteria must be used.

2.2 TESTS 1 & 2 RESULTS – VISUAL ALERTS

The objective of Tests 1 & 2 was to collect information about the visual attenson and information display to augment the existing data base. Test one investigated the variables of attenson format, display format, and pilot workload. The test provided data on the effects of a flashing master visual alert and workload on signal detection, and the effect of message format of a central alphanumeric display on alert identification performance. The test

measured the time it took the pilot to detect the "Attenson" and the central alert, the time to make a specified response to the signal, the accuracy of both the detection and response, the time to cancel the "Attenson", the sequence used in responding and a subjective evaluation of the attention-getting value of the signals.

Test two investigated the variables of visual attenson, central display location, central display cueing, and pilot workload. The test provided data on the effect of using a master "Attenson"; flashing messages, and pilot workload on signal detection. It also provided data on the effect of the location of the central alphanumeric display on signal identification. The test measured the time it took for the pilot to detect the "Attenson" and/or a central alert, the time to make a specified response, the accuracy of both the detection and the response, the time to cancel the "Attenson" or the central alert, the sequence used in responding, the performance on the flight tasks, and a subjective evaluation of the attention-getting value of the signals.

The findings of the two tests are summarized together because they both address the visual system.

The major findings of Tests 1 & 2 may be summarized as follows:

2.2.1 VISUAL MASTER

1. The results indicated significant performance degradation when a master visual attention-getter was not used. Mean detection times were substantially slower in the absence of an attention-getter in the pilot's primary field of vision. The master visual attenson located on the glareshield was most effective in gaining the pilot's attention. Neither doubling the size nor flashing the master visual attenson had a measurable effect on the pilot's detection time. Another indication of the attention-getting value of the master visual alert was revealed by a significant increase in the number of missed alerts which occurred when it was not used.

2. The data reveals that the master visual alert not only served to get the pilot's attention but also provided information on which he based response decisions. If the pilot could determine the urgency of the alert without starting his response (i.e., looking at the information display), he would respond more quickly to warnings than to cautions. This finding occurred consistently even though the mean detection times for the cautions and warnings were not measurably different.

3. The pilots preferred the steady dual visual master (master warning and master caution) to a single master visual or a flashing dual master. They felt that the steady dual master was less distracting than the other two. They also preferred the master visual attenson to the flashing box on the information display for the attention-getter of the system. but added that a combination of the two would be helpful in finding the most recent alert.

2.2.2 FORMAT

1. The format of the alerts on the central information display had an effect on response time. Cautions had a significantly shorter mean response time when the most recent caution message appeared at the top of the display rather than below the warnings. This result indicated that the pilots took more time to find the alert message when it was located in the middle of the display. This effect was counteracted by using an indication of the most recent alert (a flashing box around the message).

2. Pilots preferred to have the alert messages grouped by urgency level. They felt that this type of format would facilitate the assessment of aircraft status and result in the least important messages leaving the screen first in an overflow condition.

3. The 15° primary field of view remains the most important area for locating visual attention-getters. If there is a master visual alert in this area, the location of the information display has no measureable effect on mean detection or response times. The information display location had no effect on the number of missed alerts.

4. Pilots showed a clear preference for presenting all alert messages on the same display. They also felt that the secondary field of view (30^{o}) was an adequate location as long as there was a master visual alert.

2.3 TEST 3 RESULTS — AURAL ALERTS

The objective of Test 3 was to collect information about both verbal and non-verbal auditory alerts which could be used in the development of candidate alerting systems.

This test investigated the variables of auditory attenson, voice model, alert message format, and signal-to-noise (s/n) ratio. The test provided data on the effect of using a master "Attenson", ATC voice model, and S/N ratio on signal detection performance. The effect of message format on signal identification, the masking effects of ATC messages. and the disruptive effects of the Attenson were investigated test measurements described the time it took a pilot to detect the "Attenson" and/or the voice alert, the time to make a specified response to the signal, the accuracy of both the detection and response. the time to cancel the "Attenson" or the voice alert, the sequence used in responding. performance on the flight tasks, and a subjective evaluation of the aesthetic value of the signals.

The major findings of Test 3 - "Evaluation of Auditory Alerting System Characteristics" may be summarized as follows:

1. The test results indicate that there is a significant potential for mutual masking between synthetic speech warning messages and other voice communications within the cockpit. Performance on the ATC recognition task was substantially degraded when advisory messages were presented concurrently with voice alerts. Masking effects between competing speech sources were also evident in the response time data. It should be noted that in the testing of concurrent alert and ATC message onset, the alert messages were completely overlapped by the ATC messages; thus, it would not be appropriate to generalize these results to situations where partial overlaps will occur. The frequencies of such an occurrence in the cockpit needs to be known before a determination of the severity of this situation

can be made. These effects occurred even though the speaker configuration of the auditory alerting system was designed to optimize the pilot's ability to selectively attend to the two sources of auditory information.

2. The voice model used for the air traffic controller had no measurable effect on the pilot's ability to discriminate between alert and ATC messages and the voice model selected for programming of the Central Aural Warning System was equally intelligible with respect to male and female controller voices. Existing data on human speech intelligibility suggest that the articulation characteristics and speech rate of the individual controller, and the quality of the transmission system will be the most important factors determining the accuracy of ATC recognition. These factors are beyond the control of the alerting system designer.

3. The presence of an alerting tone as a precursor to voice warning messages did not enhance the attention-getting value of the alert in terms of time required to initiate corrective action. Some evidence suggests that response times may actually be increased slightly due to the delayed onset of the critical voice message elements when an alerting tone is used. This conclusion should be interpreted with caution in view of the particular evaluation methodology where subjects were anticipating alerts because of the abnormally high frequency with which the alerts were presented. This is a common problem encountered in attempts to simulate system malfunctions and the appropriate corrective activity that follows.

   The observed differences in response time with and without alerting tone corresponded closely to the duration of the tone.

4. The presence of an alerting tone had no effect on the frequency of ATC recognition errors or the distribution of error types. The data provides no direct evidence of increased confusion, distraction or masking effects when a voice message is preceded by an alerting tone.

5. The value of precursor tones as a source of information about the alert priority level should be given serious consideration by the alerting system designer. Individual tone signals, corresponding to distinct

levels of urgency might be used to facilitate the crew's decision-making process. In the operational flight deck environment, the frequency of alert message annunciations would be far lower than it was for this study. It may be that the attention-getting value of a precursor tone would increase as the frequency of alerts decreases. The prohibitive cost of simulation time makes the validation of this point extremely difficult. It must be concluded that the effectiveness of tones as a supplement to synthetic voice alerts depends largely on the particular application and ambient noise environment.

6. Evaluation of pilot preferences indicates that the tone-voice option was preferred by a large majority of the pilots tested. This strong preference may be due in part to pilot familiarity with alerting tones as attention-getting devices and their relatively limited exposure to the voice-only concept.

7. Findings reported in previous studies indicate that alerting tones may have important attention-getting or noise penetrating characteristics in severely degraded auditory environments. The data acquired in test 3, however, suggest that some existing guidelines and specifications requiring attention-getting tones for all cockpit voice alert messages may be unwarranted. It may be that for extremely time-critical alerts, the voice-only mode would be more effective.

8. Test results indicate that the presence of additional language context cues in the voice alert message structure provides some benefits in terms of improved intelligibility under high auditory workload conditions. Some of the gains in intelligibility due to purely redundant language may be offset, however, because of the increased time required to annunciate the essential message elements.

9. For certain types of configuration warnings and system failures, the complete sentence message format can improve voice alert effectiveness by providing more specific information about the nature and source of the problem. In some operational contexts, complete statements may be more informative than individual words or phrases.

10. The ATC recognition accuracy data do not support the hypothesis that longer voice alert messages increase the probability of interference with communications. Total error frequencies and distribution of error types remained essentially constant across voice message formats. These results are probably due, at least in part, to the fact that all voice alerts in the high auditory workload condition were presented simultaneously with ATC communications. The performance degradation associated with this treatment condition was equivalent for both word and sentence alerts. In view of this, it may be advantageous to use word-formatted messages; if the message is unintelligible when first introduced (due to a competing speech source), less time will be required before it can be annunciated the second time.

11. Although the objective performance data showed no measurable benefits for the word/phrase format, pilot preference favored the shorter message structure. Most pilots participating in the tests felt that a brief, concise statement would be easiest to identify and would tend to minimize interference with conflicting ATC communications.

12. The findings suggest that pilot information requirements and the nature of the corrective action should be the determining factors in structuring voice messages. If the flight crew is familiar with the alert message set, linguistic considerations appear to be relatively unimportant.

2.4 TEST 4 RESULTS – COMBINED VISUAL AND AURAL ALERTS

The objective of Test 4 was to collect information about the combination of visual and auditory alerts that could be used in the development of candidate alerting systems.

This test investigated the variables of alerting mode, pilot visual workload, and pilot auditory workload. The test provided data on the effect of using various combinations of tone, voice, and visual alerts, and auditory or visual workload on signal detection and identification performance. The test also provided information on the masking effects of/on ATC messages. The measurements described the time to make a specified response to the signal,

the accuracy of the response, the time to cancel the "Attenson" or voice alert, ATC recognition accuracy, the sequence used in responding, performance of the flight tasks and a subjective evaluation of the signals.

The major findings of test 4 - "Comparative Evaluation of Alerting Modes", may be summarized as follows:

1. The tracking task error measures, recorded during the post-alert segment, provided objective indices of the pilot's visual workload. Introduction of high levels of turbulence into the wind profile resulted in significant increase in deviations from the flight director target.

2. The results do not support the hypothesis that response times to alert messages would increase as a function of tracking task difficulty. Pilots were able to maintain a relatively stable level of performance on the alert identification task across low and high visual workload conditions.

3. Data indicated that the onset of alerts had a disruptive effect on the accuracy of aircraft control under high visual workload conditions. Analysis of post-alert tracking performance revealed significant increases in localizer deviations when alerts were presented by means of the combined tone-voice-visual alerting mode.

4. The simultaneous onset of a tone-voice message, master light and an alphanumeric readout seems to produce confusion. Some pilots exhibited a tendency to cross-check several sources of information prior to responding. Under these conditions, the pilot apparently finds it necessary to sacrifice some degree of control accuracy in order to identify and respond to the alert message.

5. Performance on the ATC recognition task was also resistant to degradation as a function of visual task loading. The attentional demands of the two-axis tracking task apparently did not interfere with the pilot's ability to detect and interpret ATC communications.

6. Findings reported in test 3 regarding confusion and masking effects of simultaneous voice messages were confirmed in test 4. Even though the speaker configuration of the alerting system was designed to provide an optimum environment for selective attention, the high auditory workload condition was associated with longer alert response times and substantially higher ATC recognition error rates.

7. On the basis of observed response times and total error frequencies, the tone-visual mode seems to be the most effective for the high auditory workload environment. The probability of mutual interference between alert messages and concurrent voice communications is minimized when alerts are presented by means of the tone-visual mode. Based on post-test pilot judgments, the tone-visual mode was also determined to be the preferred option in terms of attention-getting quality and overall effectiveness.

8. The distracting effects observed under high workload conditions refer only to lateral displacements relative to the flight director target. Glideslope error data showed no indication of excessive deviations during the post-alert segment. When faced with an overload condition, pilots evidently devote their attention to the control dimension that is most critical to safety of flight.

9. Contrary to the hypothesis that no single alerting mode is most effective under all combinations of environmental conditions, the data seemed to indicate that the tone-visual mode was most resistant to performance degradation under the various levels of visual and auditory task loading. This interpretation of the data should be accepted with caution in view of the simulation methodology employed. Specifically, the high auditory workload condition included no partially overlapped alert and ATC messages. Only fully concurrent messages were used which may have been a cause for excessive performance degradation. The tone-visual mode did elicit a slightly greater number of serious ATC recognition errors, although number was not significant. A flight-phase-adaptive alerting system concept may provide a mechanism to compensate for alterations in attentional demands imposed by the flight environment. This type of

system would employ computer logic or pilot action to selectively inhibit non-essential or distracting elements of the alert message as a function of flight segment or aircraft configuration.

2.5 TEST 5 RESULTS – PILOTS SUBJECTIVE EVALUATION

The objective of Test 5 was to collect subjective data on a number of alternative system concepts from current transport pilots with line experience. This test consists of two major categories of evaluations: those directed toward the visual display and those toward the auditory display.

This test was a subjective test conducted with pilots from the same corporate groups that provided the simulation data. This test investigated pilot preferences in methods of handling overflow in the central alpha-numeric display and in the number and characterization of aural alerts. The tests were single variable tests and were primarily subjective evaluation by the flight crews.

2.5.1 VISUAL DISPLAY

Three conceptual areas relating to the information display were evaluated: display format, overflow logic and separation of the alert urgency categories.

Of the three display formats that were evaluated, only two were considered as viable alternatives. The format which had all alerts entering on the top line of the display no matter what its urgency level had some support because of the ability to reconstruct the order in which faults occurred. However, the majority of the pilots preferred that the alerts be grouped within their own urgency level (the most recent alert would enter at the top of its own category). They felt that this arrangement would facilitate assessing aircraft status when there were a number of alerts present.

There were three concepts considered for accommodating those situations where there are more alert messages than can be displayed at one time. Again only two of the concepts were considered appropriate. One concept took advantage of the fact that with a display overflow it is highly probable that there are

multiple alerts within systems e.g., electrical or hydraulic. This concept would have cautions and advisories revert to a system designation. The loss of information was the biggest drawback to this concept. The second and slightly more popular alternative was to store the excess messages and provide the pilot access to the information. The pilots preferred this to be in the form of "pages" of information.

The final visual area investigated in Test 5 was the pilots ability to quickly and accurately distinguish among the alert urgency categories on the information display. All the pilots felt that it was necessary to present unique and easily distinguishable categories. They felt that color coding the caution and advisory messages was much more desirable than using the same color with a positional coding. This was especially true when a single message was present on the display.

2.5.2 AUDITORY DISPLAY

The auditory evaluations were directed toward investigating the sounds presently used in the flight deck environment to determine if they were appropriate for consideration in the advanced system concepts. The evaluations were designed to determine if there were strong stereotype responses associated with the sounds and what characteristics of the sounds were associated with the different alert urgency levels.

The results indicate that some of the sounds (bell) had very specific meanings to the pilots, while others tended to differ in meaning (low wailer), probably dependent upon previous flight experience. Since every sound had a specific meaning to some of the pilots it is recommended that the master warning and master caution tones (immediate attention) for the new alerting system be different from any sounds used in the past.

The pilots were asked to listen to the tones and classify them by urgency level. The results of this evaluation permitted the sound characteristics to be identified with alert urgency. Intermittent, wavering sounds (e.g., wailer) were selected most often as warnings. A steady insistent sound (e.g., horn) was chosen as indicating caution; the single stroke sounds (e.g., chime) were most often considered advisory.

The pilots were asked if they felt that any specific alert condition needed its own discrete alert tone given that there was a master warning and master caution tone. All the pilots had at least one specific alert that they felt needed a discrete aural. However, the only condition that was mentioned by more than 50 percent of the pilots was "Fire".

2.8 CANDIDATE SYSTEM DEVELOPMENT

The final effort in Phase I was the development of candidate alerting system concepts. Four major system components were identified; the master visual alert, the master aural alerts, the visual information display and the voice information display. The system concepts were defined as a specific set of system characteristics which utilized these major components and which met the objectives of a crew alerting system. The development effort defined two advanced systems. These systems are quite similar, differing only in the voice information display component. They both have master visual alerts for warnings and cautions. They both have master aural alerts for warnings, cautions and advisories (no other discrete aural tones are used). They both have a central visual information display that will present alert messages for warnings, cautions and advisories. The alert categories will be unique and easily distinguishable. One system will automatically present a voice alert for every warning. The second system will provide voice alerts for all warnings and cautions but only when the pilot makes an overt action (hits a switch) to request the message.

In the process of identifying the characteristics of the candidate systems, a number of questions arose. It was recognized that some aircraft conditions or situations may require crew action to be taken in an extremely short period of time. These alerts were defined as "time critical" warnings and a requirement was identified to determine the presentation media and format which was appropriate for this type of alert. These questions will be answered in Phase III of the program along with the validation of the candidate alerting system and the development of alerting system guidelines.

3.0 AIRCRAFT ALERTING SYSTEMS LITERATURE REVIEW

3.1 OVERVIEW

One of the primary ground rules for this program was that all recommendations made must be substantiated by hard data derived through empirical, subjective or analytical techniques. This section of the report summarizes the literature concerning factors that affect crew detection and response to caution and warning system alerts and displays. The information presented by this literature review will be combined with the data from the experimental and analytical portions of the program to provide the data base ffor making recommendations and developing design guidelines. The results of the experimental and analytical studies are presented in Section 5.

This literature review was structured to investigate the factors that affect crew detection and response to conventional and advanced caution and warning systems, both auditory and visual. Tactile systems are covered in an earlier report (Boucek, Veitengruber and Smith, 1977), and for that reason will not be treated here. Boucek, et. al. contains a thorough review of the literature on conventional caution and warning systems; however, little data was presented to aid designers in developing advanced auditory (voice messages) and visual (electronic displays) systems. Therefore, this section will focus on the factors that impact the design of advanced caution and warning systems: the data cited by Boucek et al will be summarized herein to enhance the self sufficiency of this report.

3.2 CAUTION AND WARNING SYSTEM FUNCTIONS

The primary functions of a caution and warning system include:

- Attracting the attention of the flight crew and directing that attention to the source of the problem. An alerting signal is required to attract the attention of the flight crew for two main reasons: (1) it is not always possible to continually monitor the status of all aircraft systems due to the many facets of flight and systems management which demand

attention, decision-making and crew response; (2) many subsystem faults are not readily apparent on the displays associated with them.

- Informing the flight crew as to the source and urgency of the problem. Sufficient information must be provided to enable them to take corrective actions and to provide feedback on the adequacy of those actions.

- Providing the crew with a mechanism(s) for exerting control over the alerting system, i.e., the termination or recall of alerts and messages.

The literature relating to system design characteristics are discussed in the following subsections.

3.3 VISUAL DATA

Visual signals are by far the most common medium used to present information in modern aircraft. The majority of visual signals take the form of small indicator lights or lighted panels with legends which illuminate when critical aircraft systems malfunction. Mechanical and electromechanical flags are also used to provide status information on the face of flight instruments when they are not working properly. Bands are used to display qualitative nonemergency conditions, i.e.. tolerance and limits on the system data displayed.

3.3.1 CONVENTIONAL VISUAL DISPLAYS

The primary signal characteristics that affect the detection of, and the response to visual signals are:

- Location
- Size
- Brightness/contrast
- Steady state or intermittent
- Color

3.3.1.1 LOCATION

The location of visual signals relative to the pilot's centerline of vision has a significant effect on not only the speed with which a signal is detected, but also the probability that it will be seen at all.

MIL-STD-411, MIL-STD-1472 and industry design guidelines presented by Van Cott and Kinkade (1972), and McCormick (1970) define the pilot's centerline of sight as a vector emanating from the pilot's eye, extending straight forward and angled 10° below horizontal. Commercial airframe manufacturers have several definitions of the centerline of sight, all of which differ from the military definition; the most consistently used definition is the line between the pilot's eye reference point and the center of his ADI.

The definitions of primary and secondary field of view also vary. The military defines primary field of view as the region within a 15° cone around the centerline of vision and the secondary field of view as the region between a 15° and a 30° cone around this centerline. Commercial aircraft manufacturers generally define primary field of view as a binocular-shaped area covering most of the pilot's primary instrument panel (containing ADI, HSI, airspeed, and altitude indicators) and secondary field of view as a binocular-shaped area covering most of the pilot's front panel (including engine instruments and autopilot mode select panels).

Human factors data indicate that most of these definitions are reasonable, however, until further testing is performed, the following criteria for location of visual alerting signals is recommended (refer to Figure 3.3.1.1-1).

- High priority alerts should be located no more than 15° from the pilot's centerline of vision.

- Lower priority signals should be located no more than 30° from the pilot's centerline of vision.

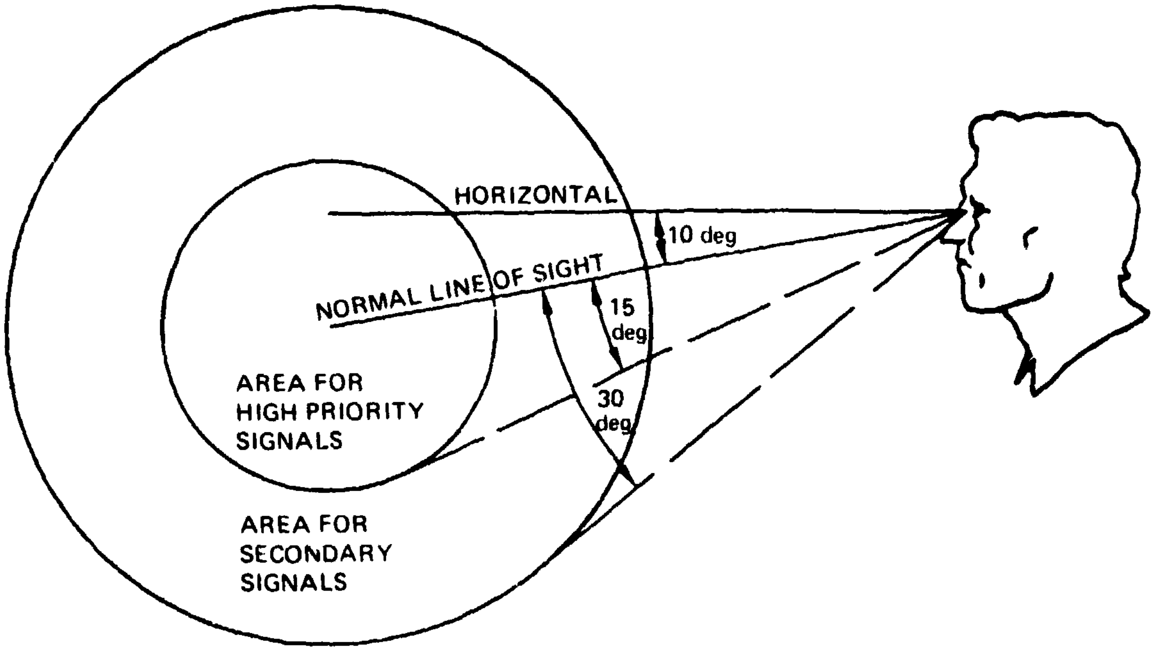


Figure 3.3.1.1-1. Preferred Placement of Visual Signals

3.3.1.2 SIZE

For visual stimuli that subtend a visual angle of 1° or less, detectability is positively related to size. However, no consistent effect of size has been demonstrated for visual stimuli larger than 1°. Sheehan (1972) measured the response times to alphanumeric legends presented on an A-7E head-up display simulator. Subjects were required to detect one of three different visual warnings (FIRE, SAM HI, or HYD PRESS), while performing a two-dimensional visual tracking task, and to respond by pushing buttons to indicate which of the three messages had been presented. The visual warnings were projected on the head-up display in one of three different alphanumeric character sizes. The character heights in degrees of visual angle and the respective reaction times were as follows: 0.5°, 1.97 second; 1°, 1.00 second; and 2°, 0.98 second. As shown in Figure 3.3.1.2-1, increasing the height of the characters from 0.5° to 1° reduced the mean response time by about one-half; however, an additional increase in height from 1° to 2° did not have a significant effect on the response time. It should be noted that the response time recorded by Sheehan included the time for detection of a message as well as the time to decide which message had been presented and to make the correct response.

Boucek, Veitengruber and Smith (1977) reported that the use of positive legend displays of proper height were found to improve response time. They found that the character height of high priority legends should be between 0.125 and 0.25 inches, whereas MIL-SPEC-18012B delineates the height-to-width ratio for warning legends to be 5:3, and the stroke width to be 0.125 to 0.166 of the height.

In summary, not much is gained when a visual signal is increased in size over 1° visual angle and there is some evidence that 0.5° is an adequate minimum. It is recommended therefore that high priority visual signals and alphanumeric legends should subtend no less than 1 degree of visual angle; lesser priority signals - no less than 0.5 degree.

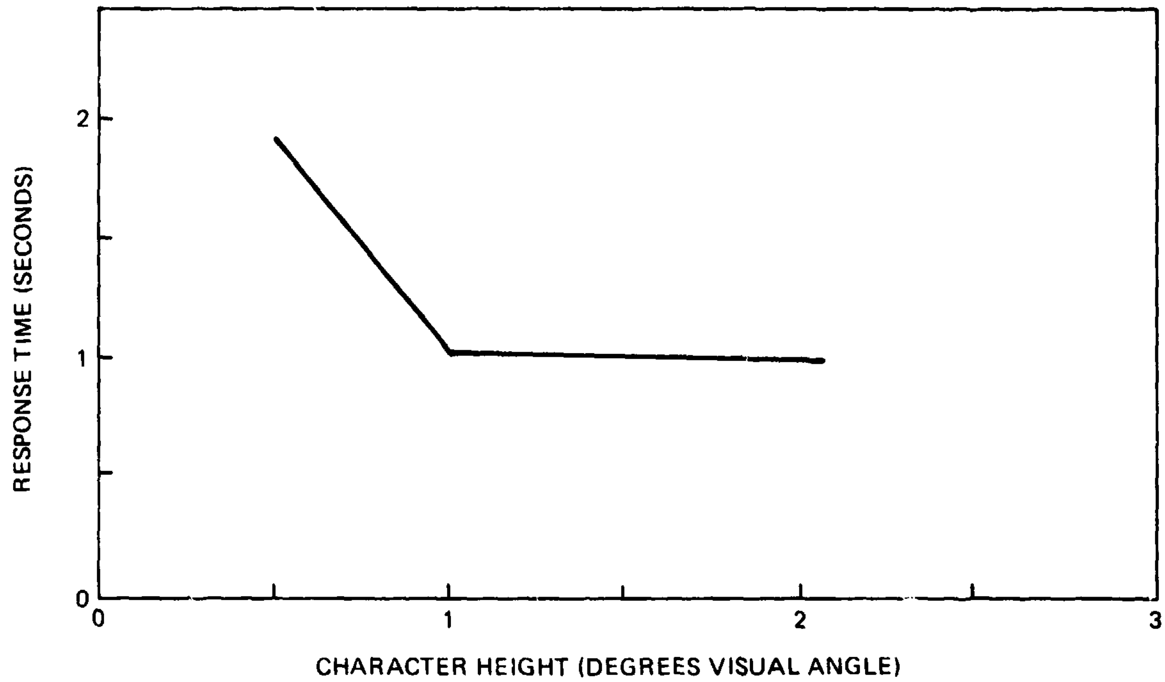


Figure 3.3.1.2-1. Effect of Character Height on Reaction Time (Sheehan, 1972)

3.3.1.3 BRIGHTNESS

The effect of signal brightness on detection is directly related to the amount of ambient lighting and the amount of light reflected by the display panel. Industry design recommendations and Military Standards give various approaches to the problem. Van Cott and Kinkade (1972) recommended that visual signals should be bright enough to stand out clearly against the panel on which they appear under all expected lighting conditions, but they should not be so bright as to impair the vision of the operator. In work stations that are darkened at night, provision should be made for dimming the warning lights when other lights are dimmed. Meister and Sullivan (1969) state that the intensity of a high-priority signal should be at least twice as bright as the immediate background; the background should be dark in contrast to the display, and should have a dull finish. Siegel and Crain (1960) state that a dark legend on an illuminated background is superior to an illuminated legend on dark background.

Although the brightness requirement of a signal is primarily determined by its criticality, the range of intensities is dictated by the detection threshold on one end, and the disruption of normal activities on the other. White and Schneyer (1960) recommended a minimum of 100 ft-L for high priority and master alerting signals, and 5 to 10 ft-L for all other signal lights. MIL-STD-411D requirements are as follows: The brightness of any rear-lighted signal shall be at least 10 percent greater than the brightness of the area around the signal. High priority signals require a recommended minimum of 150 ft-L for high ambient situations and 15 ± 3 ft-L in low ambient light. The recommended minimum brightness for secondary signals is 15 ± 3 ft-L.

In following any recommendation, care must be taken in choosing the signal values. Even though it would take a signal of 10^5 ft-L to produce actual discomfort, a direct look at a 4 ft-L signal will cause a loss of dark adaptation for a full minute (Steven, 1951).

In general, early studies agree that as signal intensity increases simple reaction time will decrease (Davis, 1947; Luckiesk, 1944; Steinman, 1944; Steinman and Venias, 1944). There is little doubt that the relationship is a

nonlinear one, and has been described by exponential, hyperbolic, and parabolic functions.

Raab and Fehrer (1962) studied the effect of flash luminance on simple reaction time using circular signals subtending 1 degree 10 minutes of visual angle and viewed binocularly in a darkened room. Using a 1-msec flash, they observed a reduction in reaction time at brightness levels out to 3000 ft-L. As shown in Figure 3.3.1.3-1, most of the improvement was seen to occur as brightness was increased to 30 ft-L, after which the reductions may be attributed to startle responses. Kohfeld (1971), used a white signal of 23^{o} visual angle, found that simple reaction time improved rapidly as brightness was increased from 0.0001 to 0.1 ft-L; improvement was less as brightness was increased to 100 ft-L (see Figure 3.3.1.3-2).

Pollack (1968) tested five luminance levels (400, 20, 1, 0.5 and 0.0025 mL) for six different colors to determine whether signal intensity had an effect on reaction time. The results of her testing agree with the previous studies.

Apparently no aircraft-related quantitative data exists defining the optimum ratio of signal brightness relative to the background. Nor has there been any data collected in an applied cockpit situation indicating how dim a signal can be before detection is impaired, or how bright it can be without blinding the pilot.

3.3.1.4 STEADY STATE OR FLASHING

A visual stimulus can be either steady state (constant brightness) or flashing (alternately on and off). Numerous experiments have been conducted on the detectability of steady and flashing lights, however, the results have been highly dependent on the procedures used by the researchers.

Gerathewohl (1953) reported that the mean reaction times to flashing lights were shorter than to steady lights of the same brightness. Crawford (1962 and 1963) found that the response to steady or flashing signal lights was affected by background conditions. Crawford's subjects were required to detect and indicate the location of signal lights. When the background was blank, either

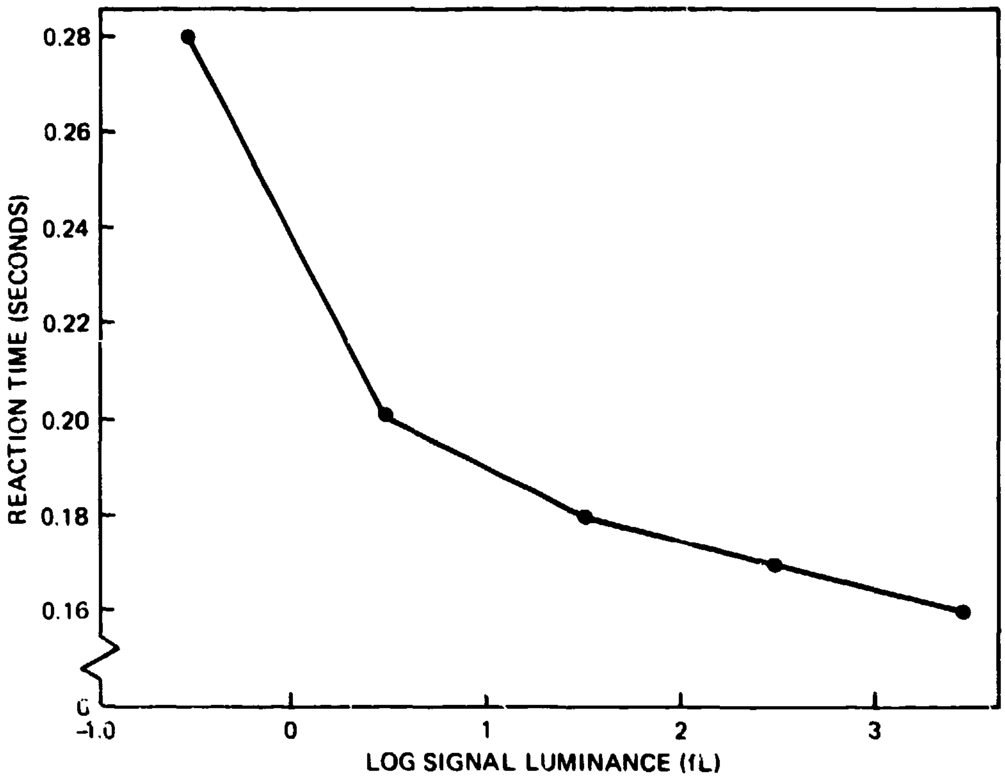


Figure 3.3.1.3-1. Simple Reaction Time as a Function of Signal Luminance (Raab and Fehrer, 1962)

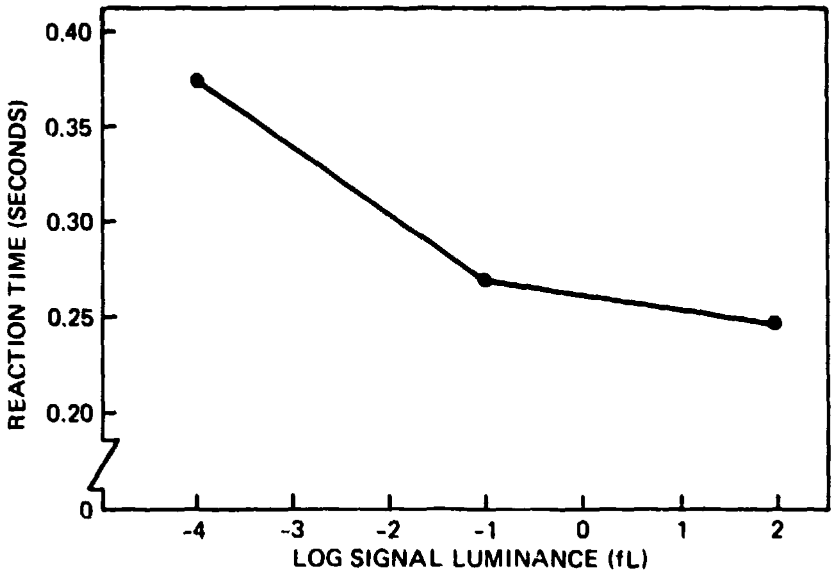


Figure 3.3.1.3-2. Simple Reaction Time as a Function of Signal Luminance (Kohfeld, 1971)

a flashing or a steady light was detected in approximately 0.8 seconds; when the background was all steady lights, flashing signal lights were detected faster than steady signal lights.

In his 1963 experiment Crawford had subjects detect steady or flashing signal lights against a background of 10 distractor lights; the number of background lights that were flashing varied from 1 to 10. The results of the 1963 experiment were similar to the results of the 1962 experiment.

It would be useful for the designer to have a method for determining which type of flashing signal is optimum for a given situation. In an attempt to provide a measure of the effectiveness of a signal in transferring information, Edwards (1971) used paired comparison techniques in which an observer had to select the most attention-getting signal from a pair. By using probability theory, he was able to construct graphic contours of equal attention-attracting power. This technique, although it has some difficulties with experimental controls, could be modified for a more realistic situation to provide a measure of the conspicuity of visual signals.

In summary, the relative detectability of flashing and steady signal lights is dependent upon whether the background lights are flashing or steady. The fastest mean detection times are obtained for flashing signal lights against a steady background; ideal visual warning system would therefore flash the warning light with all background lights steady state, either on or off.

3.3.1.5 COLOR

Numerous studies have been conducted to determine the effect of color on visual detection performance. Christ conducted a series of studies (1973-1977) investigating the performance obtained by color vs achromatic coding of visual displays. While it was concluded that color coding offers no special property, it is likely to benefit signal detection tasks. He found that color aided detection when subjects were required to deal with multiple stimulus formats, or when they were required to distinguish one class of stimulus information (signal) from another (background noise). Studies of reaction time to visual signals have indicated that color has little effect if the intensity of the signals are above 0.002 ft-L (Pollack, 1968).

Other studies (Pollack, 1968, and Haines, 1974 and 1975) showed red signals produced the slowest reaction time, while other (Coates, 1972; Weingarten, 1972) showed it to produce the fastest. Weingarten (1972) measured the relative detection times of red and green signal lights against achromatic backgrounds and found that when the background was the same luminance as the signal light, the red lights were detected 20 to 25 msec faster than the green lights. However, when the signal lights differed in luminance from the background, no statistically significant differences between the detection times of the red and green lights were found. The importance of these findings to the present study is suspect because the differences that are being discussed are in the order of 0.02 second. Therefore it can be concluded that response times to colored signals of moderate to high intensity are equal across color for dark (essentially noncolored) backgrounds.

Reynolds et al. (1972) measured the speed of detection of red, green, yellow, and white lights against copper, tan, blue and green backgrounds. The results indicate that the overall ordering of stimulus colors as measured by the speed of responding was: red, 1.8 seconds; green, 2.0 seconds; yellow, 2.3 seconds; and white, 2.7 seconds.

Hill (1974) studied the reaction time for colored signals in the whole visual field and found that no-response or missed signals may be more critical than reaction time. Figure 3.3.1.5-1 shows the percentage of no-response to blue, yellow, and green signals. A previous study by Haines (1973) also included red signals. Results were the same as the other colors up to 30° either side of center; beyond this point, the misses for the red signals increased rapidly, hitting 100 percent of the periphery of the field. Reynolds (1972) analyzed the effect of background on errors in naming the signal color. These data appear in Figure 3.3.1.5-2.

Since the results of the above experiments indicate that red signals are usually detected relatively as fast or faster than visual signals of any other color, and the current conventions (Federal Airworthiness Regulation 25.1322) dictate red signals for high-priority situations, the continued use of the following color code for cockpit signal lights is recommended:

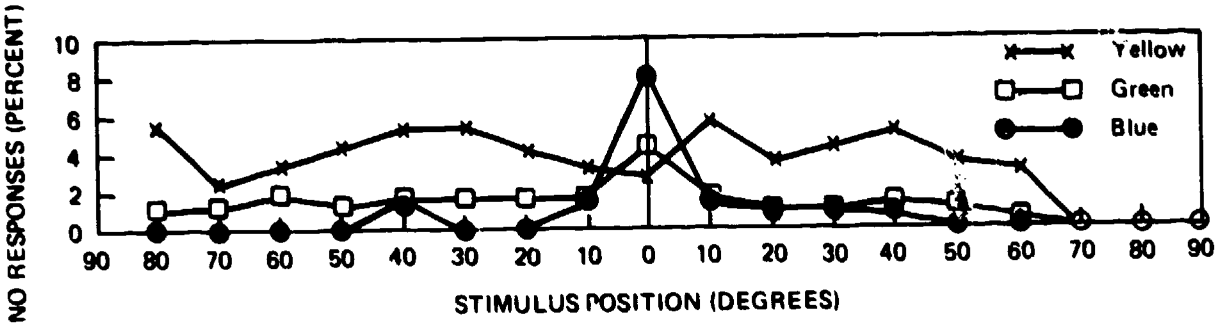


Figure 3.3.1.5-1. Percentage of No Responses to Blue, Yellow, and Green Stimuli at Equal Brightnesses Within the Binocular Visual Field (Haines, 1975)

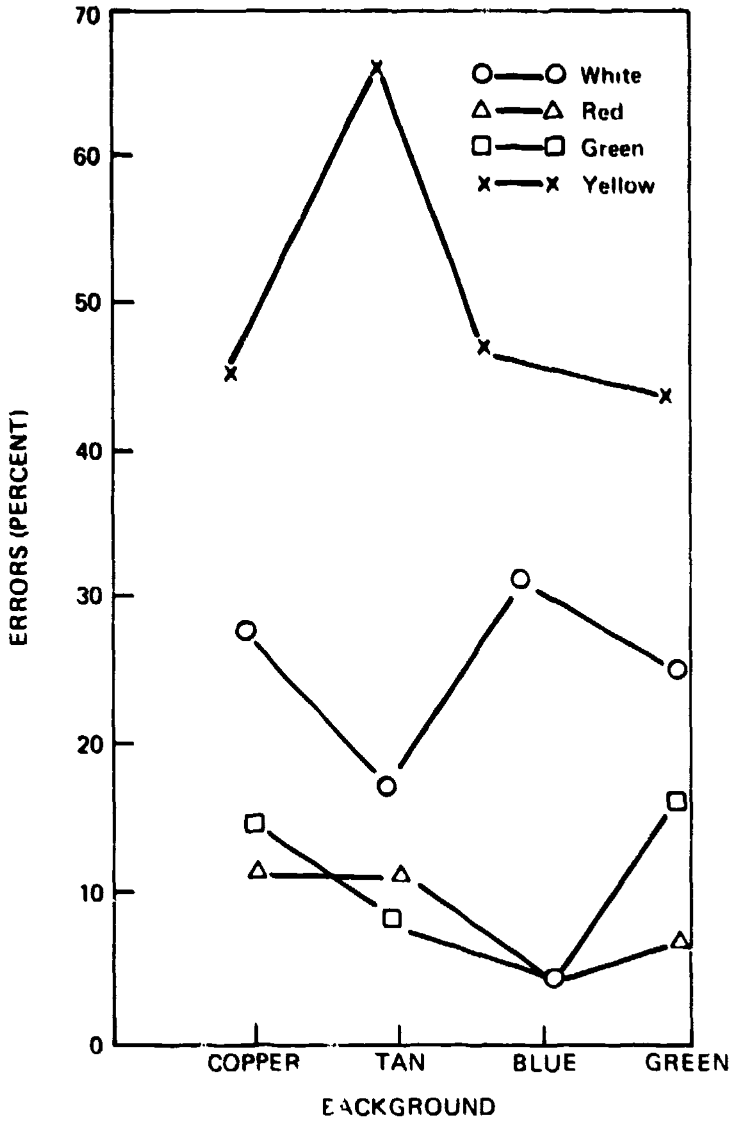


Figure 3.3.1.5-2. Interaction Between Signal Colors and Background Color-on-Color Naming Errors (Reynolds, 1972)

Red	- Highest priority warning
Amber/Yellow	- Caution
White, Green or Blue	- Normal or Safe Operation

with the color Blue being considered as a candidate for advisory alerts which have not been treated consistently in the past.

3.3.2 ADVANCED DISPLAY TECHNOLOGIES

This section summarizes the literature on display technologies that are applicable to advanced caution and warning systems. Candidate displays include:

A. Cathode Ray Tubes (CRTs)

- Shadow Mask
- Penetron
- Flat CRT's

B. Flat Panel

- Luminous - light-emitting diodes (LEDs), plasma and electro-luminescent film (ELF)
- Non-luminous - liquid crystal displays (LCDs).

Since the mid-1950's many researchers have forecast replacement of the CRT by newer "solid-state" flat-panel displays. Today, such forecasts seem even less likely to come true than they did 20 years ago. Over these years, the CRT and flat-panel displays have improved, but flat-panel display problems have proven to be more formidable and fundamental than first realized (Tannas and Goede, 1978).

The CRT, although 100 years old, is a formidable target for the challengers, and it is still improving. Since the introduction of color in the 50's, CRT's have steadily increased in brightness and luminous efficiency. At the same

time, costs have decreased due to improved manufacturing processes; life, color, performance, and most other parameters have also continued to improve.

For most applications, the CRT offers a wide range of excellent performance characteristics at relatively low cost. Nevertheless, a vast amount of research is presently being devoted to developing and improving flat-panel display technologies. In the technical thrust to develop large displays, flat panels face several hurdles in displacing the CRT; the most fundamental problems are luminous efficiency, addressing, uniformity, grey scale, color and cost.

Luminous efficiency is one of the most useful parameters in evaluating the practicality of the light emitting display technologies. Low luminous efficiency often leads to one or more of the following undesirable characteristics: high power, low brightness, high temperature, short life, bulky and expensive electronics, and a need for cooling.

A major problem with flat panel displays is the addressing and control of the many individual display elements. In a moderate-sized CRT, the equivalent of 200,000 picture elements are addressed 30 times per second at an information rate of approximately 6 MHz. No flat panel display addressing scheme has been devised to achieve comparable performance.

Duty cycle - the percentage of time available to address an individual display element - directly limits the refresh time and average brightness of a flat panel display. High power pulses can be used to improve brightness without increasing refresh time, but high voltage LSI line drivers are not yet available economically; higher currents can be used but require larger element conductors in the display itself.

The extreme sensitivity of the human eye to edge discontinuities has made the building block approach (discrete display elements) extremely difficult to implement in displays of continuous-tone imagery. The achievement of good grey-scale performance, therefore, remains a critical problem area for most flat panel displays. Pulse-width modulation and multiplexing techniques have been developed to cope with the grey-scale problem and to provide dimming

capabilities; however, at the present time such techniques are relatively exotic and expensive.

To date, no flat panel display has demonstrated good color performance at TV resolutions, however, AC plasma panels utilizing phosphor triads have come close. Several Japanese companies have recently demonstrated good color performance with panels containing only a few hundred lines.

Cost remains the most critical problem; the display electronics are usually more expensive than the panels themselves. In addition, because of the immaturity of the new flat panel technologies, production yields are at present extremely low (Tannas and Goede, 1978).

A comparison of the physical characteristics of the various displays is shown in Table 3.3.2-1. As indicated in the table, the CRT appears to offer the most advantages for implementing new integrated crew station configurations for 1980 aircraft. The advantages of the CRT include:

- High resolution
- Good addressability
- High contrast
- Flexibility
- Color capability
- Relatively low cost.

However, the CRT still has several disadvantages:

- High voltage requirements
- Large volume
- Extra circuitry required to achieve corner edge focus
- Limited useful life under high ambient light conditions
- Implosion hazard.

For these reasons it may be desirable for aircraft applications of the 1990's to replace CRT's with flat panel displays (Hatfield, Robertson, Batson, 1979).

Flat panel technologies have shown much promise, with attendant advantages such as:

Table 3.3.2-1. Comparison of CRT and Flat Panel Capabilities

Characteristic	CRT	Ac plasma	Dc plasma	Ac elf	LED	LCD
Size						
Diagonal (in)	To 25	To 16	To 16	6	5	1.4
Depth (in)	10 to 12	3		1.5	1	1
Resolution (typical) (lpi)	100	50	50	68	64	100
Contrast ratio	10:1	25:1	25:1	25:1	>25:1	30:1
Grey scale	16	Bilevel	Bilevel	16	10	10
Color	Full	Orange	Red, green, blue	Yellow, orange	Red, green, yellow	Mono
Life (hr)	10,000	> 10,000	> 10,000	> 10,000	12,000	> 20,000
Interface complexity	3	Row and column connections				
Advantages	Resolution Full color Display area	Rugged Power Depth Memory	Memory		Color	
Disadvantages	Hazards Depth Not rugged Power	Interface Display area				

- Direct digital compatibility
- Inherent accuracy and edge focus
- Small area and large area contrast
- Relatively low supply voltages
- Very small volume.

Flat panel plasma displays are considered the most formidable contenders to CRTs in the near future. Plasma displays easily fit spaces too small for CRTs. They also have high readability, wide viewing angle (160°), and are suitable for graphic and alphanumeric display applications. Typical brightness is 30 ft-L with contrast ratios of 25:1. A plasma panel's life is longer than a CRT's (greater than 10,000 hours) but is available primarily in only one color - orange (Hassberg, 1979).

3.4 AUDITORY DATA

Auditory warnings can be used to alert personnel to impending dangerous conditions, to critical changes in system or equipment status, and to critical actions that must be taken. The detection of auditory warnings is affected by the characteristics of the signal, the inherent capabilities and limitations of the individual listener, and the listening environment. This section describes the parameters that impact the detection of, and the responses to, alerting tones and voice messages.

3.4.1 ALERTING TONES

The primary factors that impact the detection and response to nonverbal auditory alerts are:

- Frequency
- Intensity
- Location of Sound Source
- Message Content
- Number of Tones
- Environmental Factors

3.4.1.1 FREQUENCY

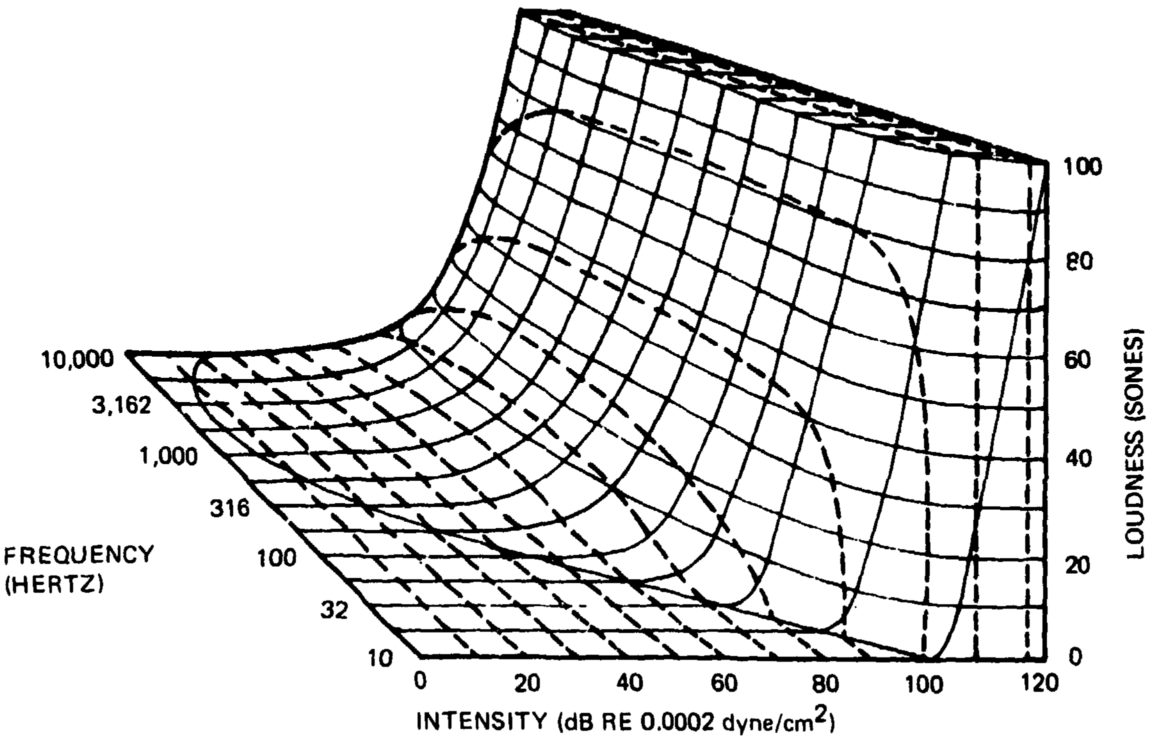
Frequency has a strong effect on perceived loudness. Young humans can detect sounds with frequencies ranging from 20 Hz to about 20,000 Hz. Midfrequency tones (2000 to 4000 Hz) tend to sound louder than lower or higher frequency tones of the same energy (Fletcher and Munson, 1933), however, this effect of frequency on perceived loudness decreases as sound amplitude increases.

Two additional frequency-related factors that impact the detection of aural signals are aging, which causes a progressive loss of hearing at higher frequencies and ear injuries (Morgan, 1943). For these reasons, it is important that signaling devices do not use a single frequency, but rather a combination of two or more widely spaced frequencies between 250 and 4000 Hz (Boucek, Veitengruber and Smith, 1977).

3.4.1.2 INTENSITY

Before summarizing the literature on the impact of intensity on detection, the difference between intensity and loudness must be clearly understood. Intensity is a physical measure of the energy level of a sound transmitted through a unit of area. Loudness, on the other hand, is an attribute of the sound as heard and reacted to by a listener. Loudness is therefore a subjective response that depends primarily on intensity, but is also affected by frequency, as indicated earlier. The relationship between frequency and the two dimensions of sound, i.e, loudness and intensity, is shown in Figure 3.4.1.2-1.

As a general rule, an intense sound is more likely to be detected than a quieter sound of the same frequency. However, the detectability of any particular sound is primarily dependent on background noise. For any given background conditions, there is an intensity of signal that will be detected 50% of the time by a particular individual; this level of intensity is referred to as the threshold intensity. An increase in intensity of as little as 3 dB over threshold can result in nearly 100% detection.



Note: Subjective loudness in sones is represented vertically above the intensity-frequency plane. The heavy curves coursing from front to rear in the diagram are equal-loudness contours for pure tones. (Stevens and Davis, 1938).

Figure 3.4.1.2-1. Three-Dimensional Surface Showing Loudness as a Function of Intensity and Frequency

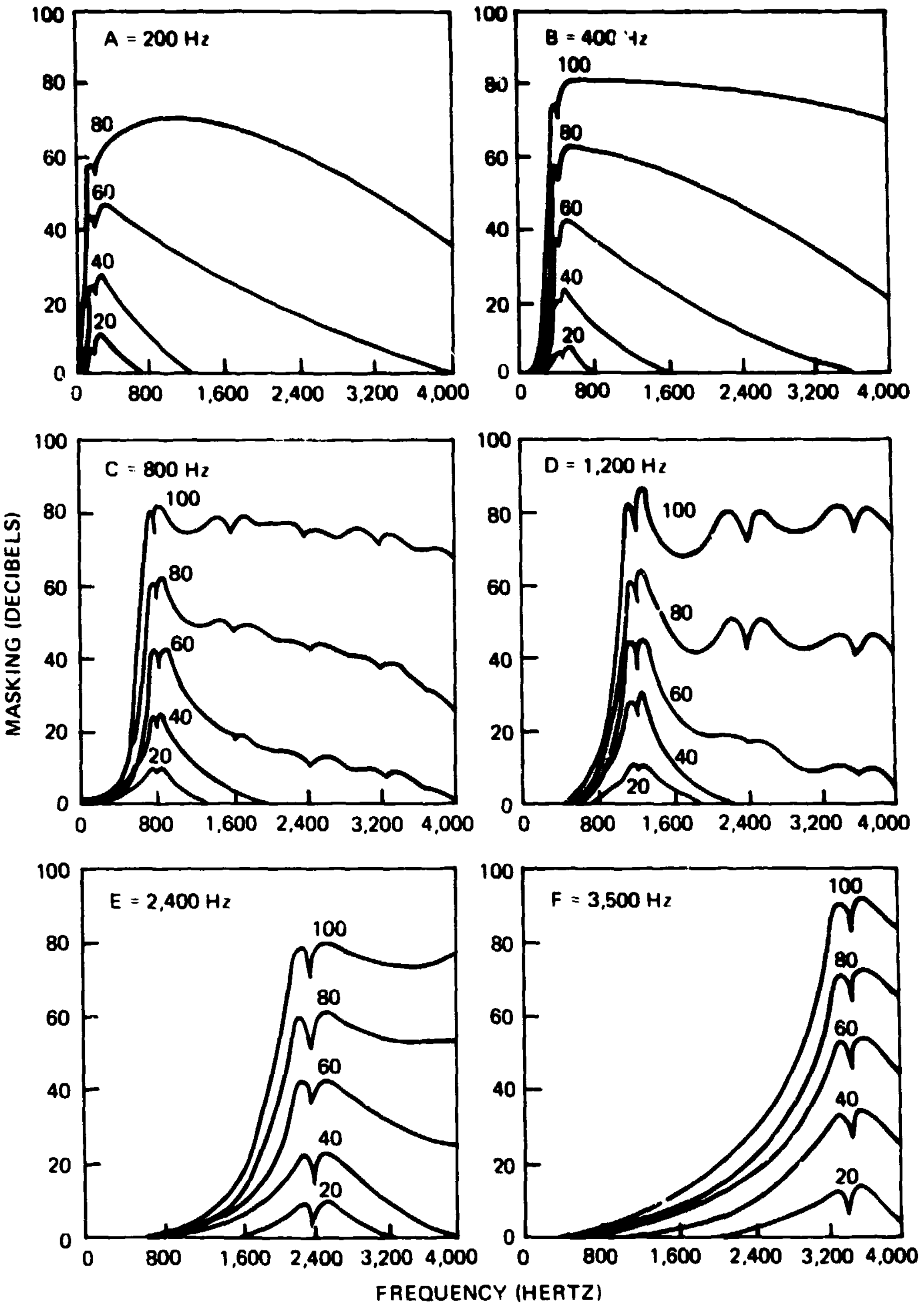
Noise mixed with a signal tends to raise the detection threshold above the "threshold in quiet"; this effect is referred to as masking. Since auditory alerts will be used in an environment where the background noise is constantly changing in amplitude and frequency, it is important to determine what aspects of the background noise will require adjustments in signal intensity. For cockpit applications the masking of aural alerts should be evaluated for the following three types of ambient noise.

Noise Type	Distinguishing Characteristics
Pure tone	Bandwidth = nominal frequency ±0 Hz
Narrow-band noise	Bandwidth = nominal frequency ±45 Hz
Wide-band noise	Bandwidth = wide spectrum

Quantitative relationships between the frequency of the masking tone and the amount of masking of auditory signals of various frequencies are shown in Figures 3.4.1.2-2, -3 and -4. In Figure 3.4.1.2-2, the frequency of the auditory signals (masked tones) are given on the abscissa of each graph. The ordinate presents the level, i.e., the amount above the threshold-in-quiet level, that the auditory signal must be elevated in the presence of the masking tone.

The number on each represents the intensity of the masking tone, measured as the amount above the threshold-in-quiet level. The lowest curve in Figure 3-9 gives the threshold-in-quiet values.
Maximum masking of a pure tone occurs when the background sound is of the same frequency range as the signal. Substantial masking also occurs when the auditory signal is composed of higher frequencies than the ambient noise; lower frequency alerting signals are significantly less subject to masking. The masking effect of narrow-band ambient noise is similar to that described for a pure-tone environment.

For cockpit applications, wide-band noise effects must also be considered. Morgan, et.al. (1963) state that the masking effects of wide-band ambient noise are considerably different than the masking effects of narrow-band and pure tone ambient noise; the effects of wide-band noise extend beyond the



Note: Number at top of each graph is frequency of masking tone, and number on each curve is level above threshold of masking tone.

Figure 3.4.1.2-2. Masking of One Tone by Another Tone (Wegel and Lane, 1924)

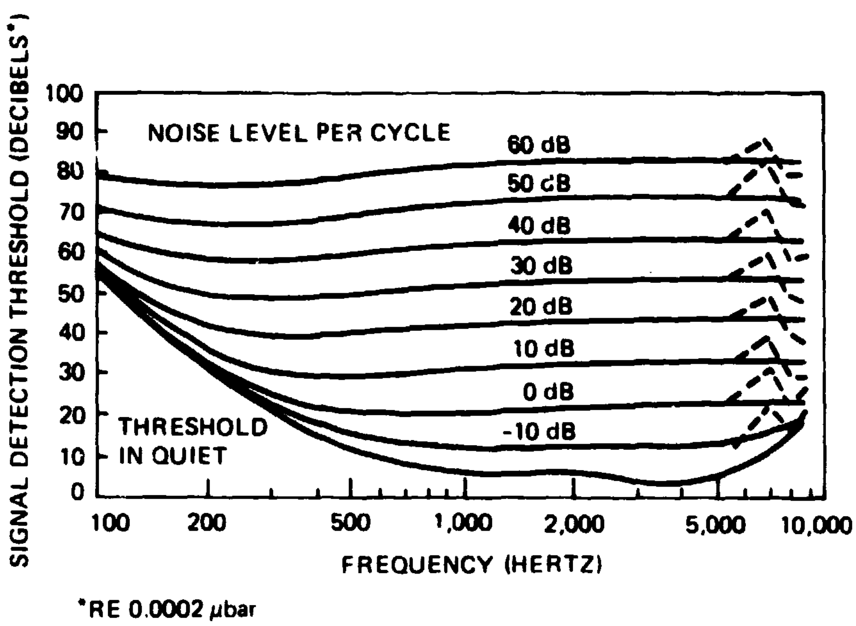


Figure 3.4.1.2-3. Masking Effect of White Noise on a Pure Tone (Hawkins and Stevens, 1950)

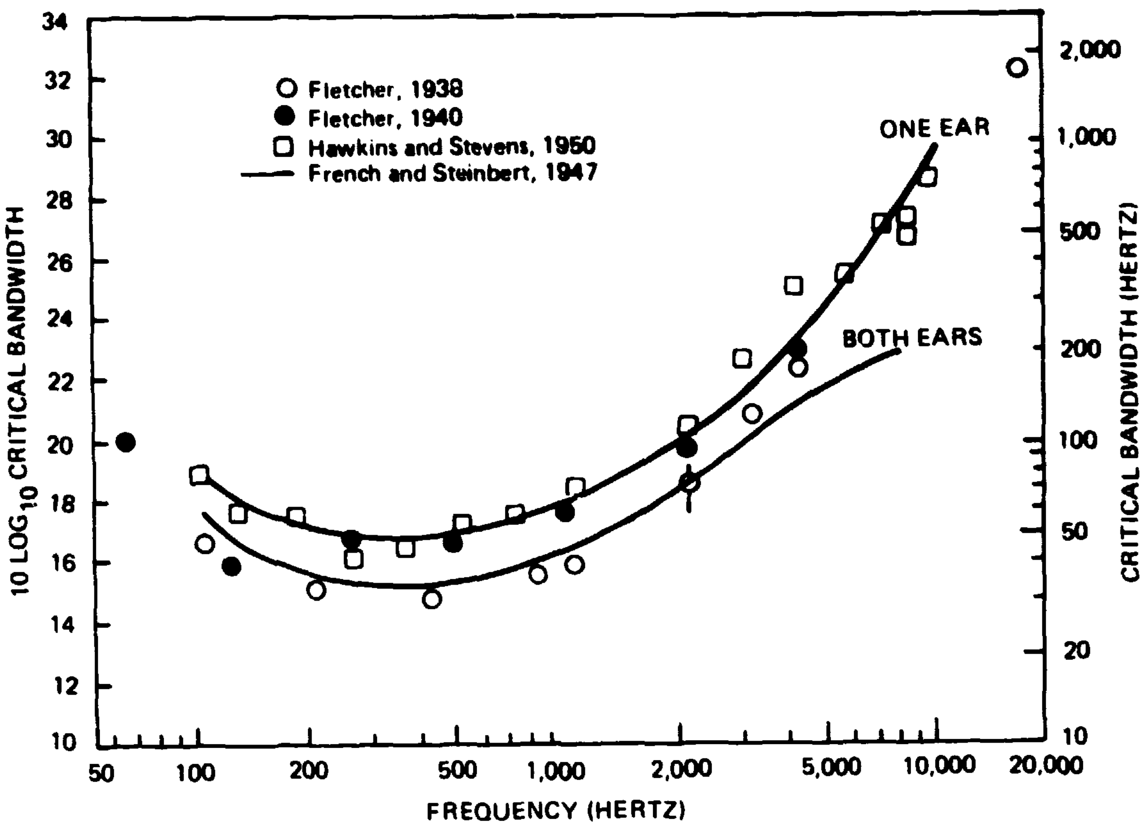


Figure 3.4.1.2-4. Critical Bandwidth of Masking in Wideband Noise (Fletcher, 1953)

spectrum of the noise itself. The masking effect of wide-band noise that has the same intensity throughout the spectrum (white noise) is approximately linear with respect to the increase in intensity of the noise. For wide-band noise that does not have uniform intensity over the frequency spectrum, the ear as the ability to filter or reject the part of the noise that is outside a certain range around the signal, thus eliminating some of the noise and making the signal more audible. The width (in Hz) of this range is called the "critical bandwidth" and varies with the frequency of the signal being used. Morgan, et.al. (1963) state that the threshold of a pure-tone aural alerting signal can be predicted if the spectrum of the noise near the frequency of the tone is known. In making this prediction, it is assumed that the masking is being done by the noise components near the frequency of the signal, those that lie in the critical bandwidth. When used to predict masking the critical bandwidth is defined so that the sound pressure level of the noise in the critical band is equal to the sound pressure level of the signal at its masked threshold (the intensity where 50 percent of the signals are detected when noise is present). Morgan presented the following procedures for predicting the masked threshold of an aural alert signal at any signal frequency in wide-band noise:

1. Measure the level of the ambient noise at the auditory signal's frequency.

2. Correct this measure level for the wide-band effect by adding the 10 $\log_{10}$ value of the critical bandwidth (read directly from the left ordinate in Figure 3.4.1.2-4).
3. This corrected value is the masked threshold of the aural alert.

These methods are directed toward detecting pure-tone signals, which are harder to detect in noise than multifrequency signals. Van Cott and Kinkade (1972) presented two well-accepted guidelines for multifrequency auditory signals:

1. A sound signal should exceed its masked threshold by at least 15 dB for good discrimination.

2. An optimum signal level in noise is halfway between the masked threshold and 110 dB.

Also to be considered when working with any type of aural alerting signal is MIL-STD-1427B, which requires that auditory signals have a signal-to-noise ratio of a least 20 dB.

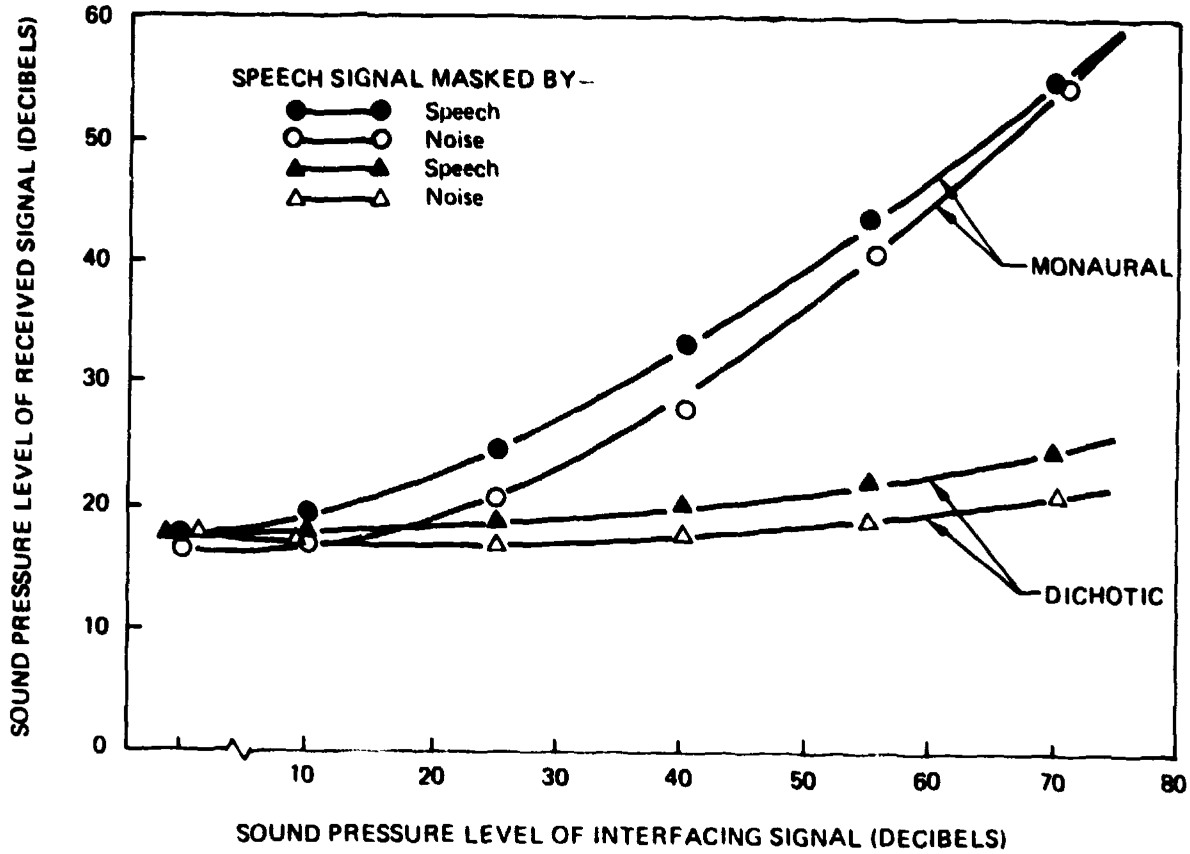
However, these guidelines may in some cases conflict with the pilot criticism that most aural alerts as currently implemented are too loud (Cooper, 1977). Care must be taken when applying these guidelines to the actual cockpit environment because it is possible to introduce sound levels that are intolerable to the pilot. The range of signal intensity by necessity must be limited on one end by the auditory threshold and at the other end by the onset of pain (110 dB). The intensity/exposure time interaction, which imposes limits after which there is a high risk of damage for unprotected ears must also be considered. Thus, until data that resolves this conflict are obtained, it is recommended that the following guideline be used:

	Threshold + 15 dB	
Signal Loudness =	or	Whichever is less
	Threshold + 1/2 (110 dB-Threshold)	

3.4.1.3 LOCATION

The masking effects of background sounds are also dependent upon their location relative to the signal sounds; signals perceived as coming from a separate location are more likely to be detected.

Egan, Carterette, and Thwing (1954) had subjects listen to messages under monaural and dichotic conditions. In monaural listening, the message to be received and interfering noise or messages are both received in the same ear, in dichotic the message is presented by an earphone to one ear, and interfering noise or messages are presented by another earphone to the other ear. Dichotic listening was found to provide location cues which enabled the subjects to discriminate between signals and noise. As can be seen in Figure 3.4.1.3-1 dichotic listening provided an advantage equivalent to an increase



Note: Curves show threshold sound pressure level for perception of a received signal masked by an interfering signal.

Figure 3.4.1.3-1. Comparison of Dichotic and Monaural Masking (Egan et al., 1954)

of up to 30 dB in the intensity of the signal message. For this reason, MIL-STD-1472B recommends that dichotic presentation be used, alternating the source of the signal from one ear to the other via a dual channel headset.

The advantages offered by dichotic earphone presentations are not to be expected in an open (non-headphone)environment. To approximate an open type of situation, Speith, Curtis, and Webster (1954) tested subjects by presenting simultaneous messages from the same (single-source condition) or from two different loudspeakers that could be separated from each other horizontally. When both messages were transmitted from the same loudspeaker, the subjects responded correctly 66% of the time. Response improved to 86% correct for 10^0 to 20^0 separation of messages, and to 92% for 90^0 to 180^0 separation. Speith et al., did not determine how much increase in signal message intensity would produce the same amount of improvement produced by the separate conditions. The ability to localize a signal is also dependent on the frequency of the sounds. Mills (1958) found that localization of pure-tones was optimum over the frequency ranges of 250 - 1000 Hz and 3000 - 6000 Hz. Localization was found to be poor over the range of 1000 - 1500 Hz, and for frequencies around 8000 Hz. It was also found that broad-band signals are generally localized much better than pure tones. It can be concluded that for binaural listening, broad-band signals are more likely to be detected than pure-tone signals.

Cherry (1953) addressed the problem of how a critical verbal message is detected when other messages are occurring at the same time. He presented subjects with two messages, either mixed to both ears, or one to the left ear and the other to the right ear, and asked them to repeat one of the messages. It was found that some messages could be separated if they were presented in the mixed fashion and some could not. However, the subjects had no trouble separating the messages when they were presented to different ears. In fact, after the subject was comfortably repeating one of the two messages, the message in the other ear was switched to German; no observer detected the switch.

The detectability of a verbal message is often affected by the content of the message itself. For example, a person's name is usually more attention-getting than any other auditory message of the same volume (Howarth

and Ellis, 1961). Experimental data indicate that having a person's name precede an auditory message appears to have about the same effect on detection as increasing the loudness of the message by 3 dB. Thus, it is recommended that aural alerting messages be preceded by an identifier to which the pilot is more than normally sensitive, e.g., the pilot's name or aircraft identification.

3.4.1.4 MESSAGE CONTENT

The detectability of nonverbal auditory signals is also affected by content. Keuss (1972) used two signals in close succession; the first signal (essentially a ready signal) was given for 25 milliseconds and then the response signal was presented; the test subjects were required to acknowledge when they heard the second signal. Generally, it was found that reaction time varied inversely with the intensity of both signals; however, as signal intensities were increased to 110 dB, reaction time increased, probably due to the startle effect.

Siegel and Crain (1960) ran an experiment under night conditions. Subjects were required to perform a tracking task and respond to a warning signal which was either a light, a single tone, or a double tone. Use of the two-tone auditory signal resulted in significantly shorter response time.

The auditory sense adapts rapidly to constant stimulation; steady-state signals tend to become less noticeable after a short period of time. Therefore, a steady-state signal that is not detected immediately is likely to go unnoticed. The auditory system does not adapt as rapidly to intermittent or changing signals as it does to steady-state signals, and for this reason intermittent signals are more likely to be detected than steady-state signals. MIL-STD-411D requires that an auditory master warning signal have an 0.85 second ON time and an 0.15 OFF time, with the cycle continuing until the system is de-energized.

3.4.1.5 NUMBER OF TONES

Experimental testing has shown that humans can make precise judgments about minute differences between stimuli, however, they are limited in their ability to make absolute judgments (Miller, 1956). In other words, when presented with two signals, a person can tell quite accurately whether they are different. Shower and Biddulph (1931) reported that under ideal conditions, listeners could detect frequency differences between tones as small as 2 or 3 Hz. Pollack (1952) had listeners identify tones of different frequencies by assigning numbers to them. When only two or three tones were used the listener never confused them, but with five or more tones confusions were frequent. MIL-STD-1472B states that "if absolute discrimination is required the number of signals shall not exceed four". Cooper (1977) surveyed major airplane manufacturers and found that they recommended that the number of audio warnings or signals should be limited to four or five. Just recently the S-7 committee of the SAE issued a draft version of a revision to Aerospace Recommended Practice - 450C which reduced the recommended number of discrete aural warnings, other than the master warning, caution and advisory tones, permitted in the flight station.

Pollack and Ficks (1954) found that using sounds which differed in more than one dimension (e.g., frequency, purity, etc.) increased the total amount of information transmitted; however, the amount of information conveyed would decrease as more dimensions were used. This data leads to the conclusion that if more than four or five auditory signals are to be presented, they should differ in more than one dimension.

3.4.1.6 ENVIRONMENTAL FACTORS

Distractions, existing cognitive workload, and the vigilant state of the operator also have an effect on response. In general, any environmental factor that increases the demands on an individual will decrease the likelihood of signal detection and increase response time. Many studies have demonstrated the impact of distracting stimuli on signal detection and response. Adams and Chambers (1962) found that the addition of irrelevant auditory distractors producted a detriment in performance on a visual tracking

task. Likewise, irrelevant visual distractors degraded the performance of an auditory tracking task. One recurring findings in these studies is that the detectability of stimuli, in the presence of distracting signals, is enhanced by bimodal signal presentation. Buckner and McGrath (1961) had subjects perform a vigilance task in which they were presented 24 signals during a 60-minute session. For any one session, all of the signals were either visual, auditory or combined visual and auditory. The detection rate for all three types of signals was close to 100% at the beginning of the sessions, but decreased with time; however, the detection rate was higher for the bimodal signals (89%) than either unimodal signal (auditory 84%, and visual 72%).

The higher the workload the more likely that a signal (especially visual) will go unnoticed (Boucek, Veitengruber and Smith, 1977). To increase the likelihood of detections, a warning signal should change the sensory environment (Senders, 1952). Also the probability of an observer detecting a particular signal changes considerably with time, even when the signal and environment are constant; these changes in observer efficiency are usually ascribed to changes in the observer's vigilant state. Also, low signal presentation rates have a detrimental effect on their detection (Adams, Humes and Stension, 1962). For these reasons it is recommended that bimodal signals be used to present high-priority alerts.

3.4.2 VOICE WARNING SYSTEMS

3.4.2.1 OVERVIEW

The primary functions of a caution and warning system are to alert the flight crew of a problem, and to inform them of its nature. It should also guide crew action, and provide feedback on the status of the system after the crew has responded. These functions have been typically handled by a combination of visual and auditory techniques. Auditory signals, in particular, have become so numerous that it is quite difficult to associate the various sounds in use with specific failures; Veitengruber, et al (1977) report that as many as 17 distinct aural alerts are used on present wide body aircraft. This problem of associating various aural signals with a particular failure is eliminated if verbal messages are used.

Under high-stress conditions, the pilot's audio-visual load may reach saturation, causing a potential decline in his efficiency and performance. Therefore, a system that can convey warning information under these conditions without further degrading performance is essential. One way to accomplish this is to provide more information per message and to allow the transmission of only absolutely essential messages. A warning system using voice messages to inform the pilot of aircraft status and incorporating a priority attenuation system meets these criteria. Research indicates that the use of verbal warnings produces significant improvement in response time, especially during periods of heavy workload or stress (Pollack, et. al., 1958; and Kemmerling, et. al., 1969). Another important advantage of the voice warning system is that is allows the pilot to evaluate the criticality of a situation without bringing his eye scan back into the cockpit. Given that a voice warning is to be used the factors that impact its effectiveness have to be optimized. These factors are described in the following paragraphs.

3.4.2.2 SPEECH GENERATING TECHNIQUE

There are two basic types of voice systems. Voice-modeling uses prerecorded presentations of actual male or female speech for every possible caution and warning; resulting messages are very close to every-day speech and relatively easy to understand. The other approach uses a digital voice synthesizer to generate messages by combining distinct phonemes. The major advantage of this technique is the reduction in size and weight of the equipment required; a drawback is that synthesized voice is not realistic since much intonation is missing, and some sounds are difficult for the synthesizer to reproduce.

Kerce (1979) conducted a study that compared the intelligibility of the two speech generating techniques and found that greater intelligibility results from the use of voice-modeling; the test participants also expressed a strong subjective preference for voice-modeling.

3.4.2.3 ALERTING TONE REQUIREMENTS

MIL-STD-1472B states that verbal warning messages shall consist of an initial attention-getting tone, and a brief standardized speech message to identify

the problem and suggest appropriate action. MIL-STD-1472B also states that where reaction time is critical, the initial alerting tone should be of 0.5 second duration, and that all essential information shall be transmitted in the first 2.0 seconds of the identifying or action signal.

Douglas Aircraft administered a questionnaire to 131 flight operations personnel; 71% of the respondents agreed that voice warnings should be preceded by an alerting tone.

Whereas military standards and subjective pilot preferences favor a precursor tone, empirical data fails to support its inclusion in a voice warning system. Simpson and Williams (1978) conducted a study to determine the impact of a precursor tone on pilot detection and response times. The results indicated that the inclusion of the alerting tone added about one second to the duration of the warning signal, and resulted in a 0.65 second increase in response time. They concluded that the tone seemed to be superfluous, since it takes extra time to present, and does not facilitate pilot detection or response times; they also stated that the difference in voice quality between normal speech in the cockpit and the voice warning system may itself be sufficient to direct attention to the message.

3.4.2.4 VOICE MODEL

The major factor in selecting a voice model (either male, female or electronic) is its intelligibility in the cockpit. Douglas Aircraft has administered several questionnaires to pilots (Erickson, 1978; and Kerce, 1979) which solicited preferences for male or female voice models; the results of these surveys indicated a distinct preference for female voice messages. Several of the pilots indicated that they preferred the female voice because most of the communication in today's cockpit is normally delivered by male voices. Simpson and Williams (1978) state that the voice model selected for cautions and warnings should be qualitatively different from the other voices in the cockpit, and that since both male and female speech are present in today's cockpits, voice warnings should be given with synthetic speech.

Very limited objective performance data exists on the relative intelligibility of male, female and synthetic voice models. Kerce (1979) conducted two studies that compared the relative intelligibility of the three voice models. The results indicated that the female voice was found to be more intelligible than the male voice and that both were more intelligible than the synthetic voice (see Figure 3.4.2.4-1). However, it should not be inferred that similar results would be obtained in any male-female voice comparison, as intelligibility is greatly affected by the voice models and recording procedures employed.

3.4.2.5 VOICE INFLECTION

Very limited data exists to aid designers in selecting between conversational, monotone or urgent voice inflections. Two unpublished studies conducted by Douglas Aircraft found that pilots preferred the monotone inflection. In another subjective study conducted by Douglas Aircraft (Erickson, 1978) 60% of the airline personnel suggested that voice inflection should vary with the nature of the alert, either advisory, caution or warning. In the end, Douglas selected a monotone inflection for their DC 9-80 voice warning system; they felt that the distracting effects of urgent-sounding messages outweighed any potential benefits.

3.4.2.6 MESSAGE FORMAT

The wording of cockpit voice messages has a significant impact on the effectiveness of an aural caution and warning system. Little empirical data exists on message length, content, phonetic confusability, and message structure, however, for consistency the structure and content of the aural messages should be similar to the messages presented by the visual indicators or the central display.

Highly discriminable keywords or phrases should be used such that the messages are easily understood with as little demand on pilot workload and attention as possible (Simpson, 1976). Several studies have been conducted to determine whether additional syllables or words, or a sentence context would increase intelligibility. The results of these studies indicate that linguistic

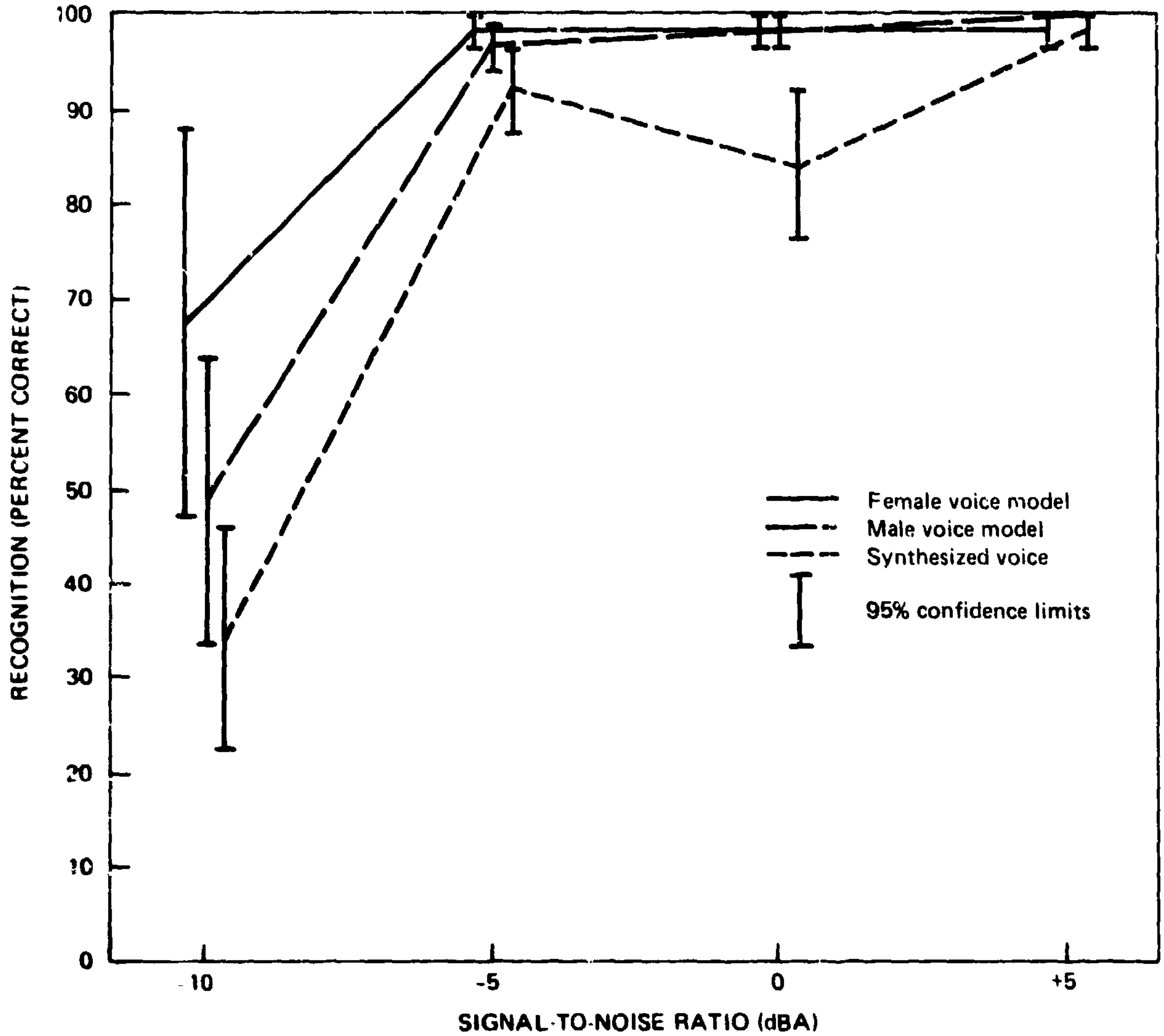


Figure 3.4.2.4-1. Intelligibility as a Function of Voice Quality and Signal-to-Noise Ratio With Competing Speech Noise Background (Kerce, 1979)

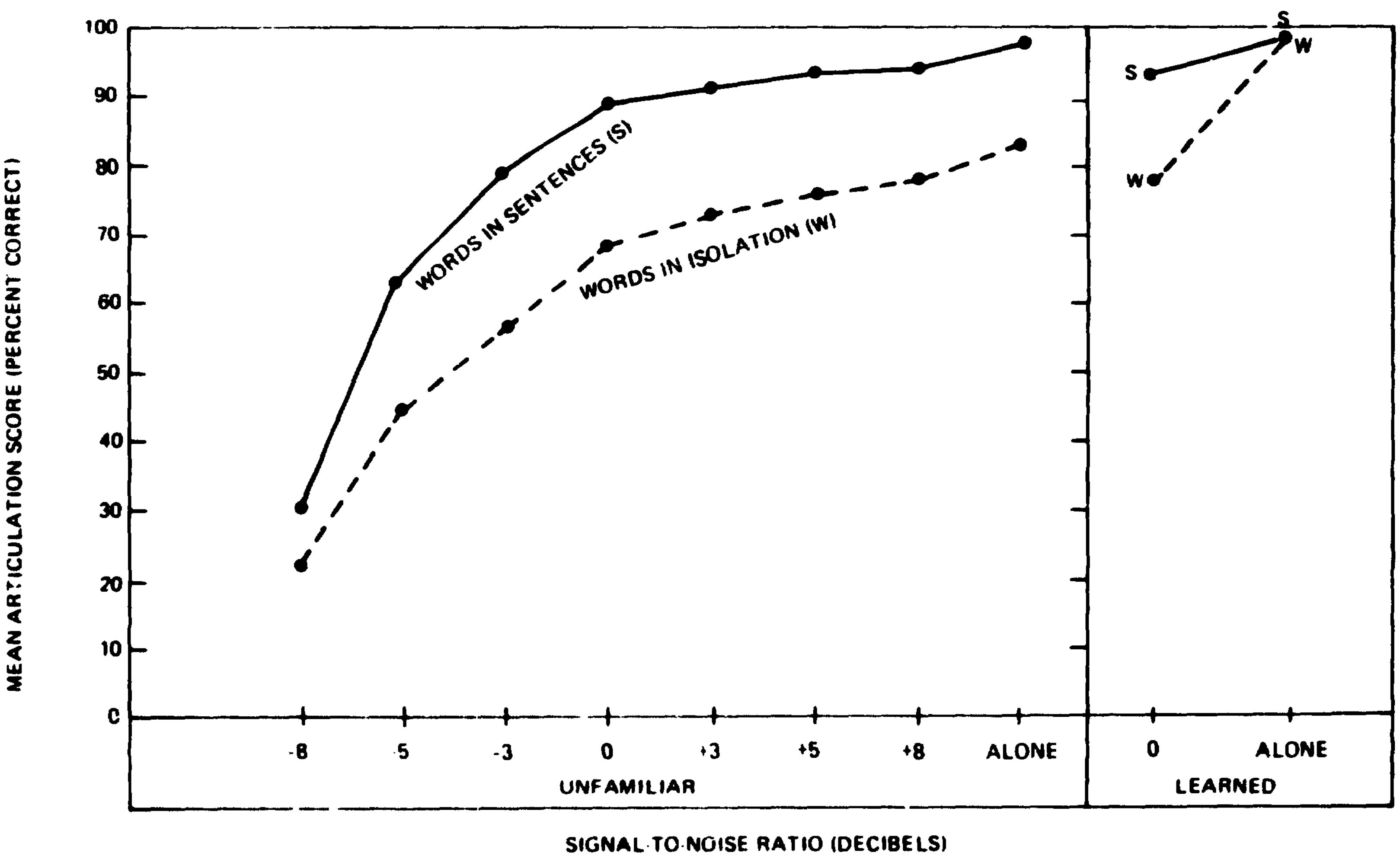
redundancy facilitates the comprehension of aural messages and tends to reduce response times. Hart and Simpson (1976) found that aural messages presented in a sentence format were more intelligible than two-word messages and required fewer repetitions for comprehension. Simpson (1976) presented synthesized mono- and polysyllabic keywords and sentence-length messages to airline pilots under several signal-to-noise ratios. She found that sentence messages consisting of monosyllabic keywords were responded to more accurately, over a wider range of signal-to-noise ratios. (Figure 3.4.2.6-1). Polysyllabic words did not show this tendency; scores for both sentences and isolated words were approximately the same (Figure 3.4.2.6-2). These data seem to indicate that pilots need some "warmup" or alert to enable them to receive the verbal message rapidly and accurately; the monosyllabic words did not give the pilot enough time to prepare for message reception. Response time results for these data are shown in Figure 3.4.2.6-3.

The advantages of linguistic redundancy may disappear if an alerting tone or word precedes the voice message; the use of single keywords or phrases might therefore prove to be as good as or better than the sentence format. Brevity is highly valued for cockpit communications because it is believed to reduce the time required for comprehension and the probability of overlap with other auditory events; also, brief wording conforms to the current style of cockpit communications. Finally, each additional word of digitized speech represents additional memory and associated costs (Simpson, 1976).

3.4.2.7 INTENSITY AND CONTROL

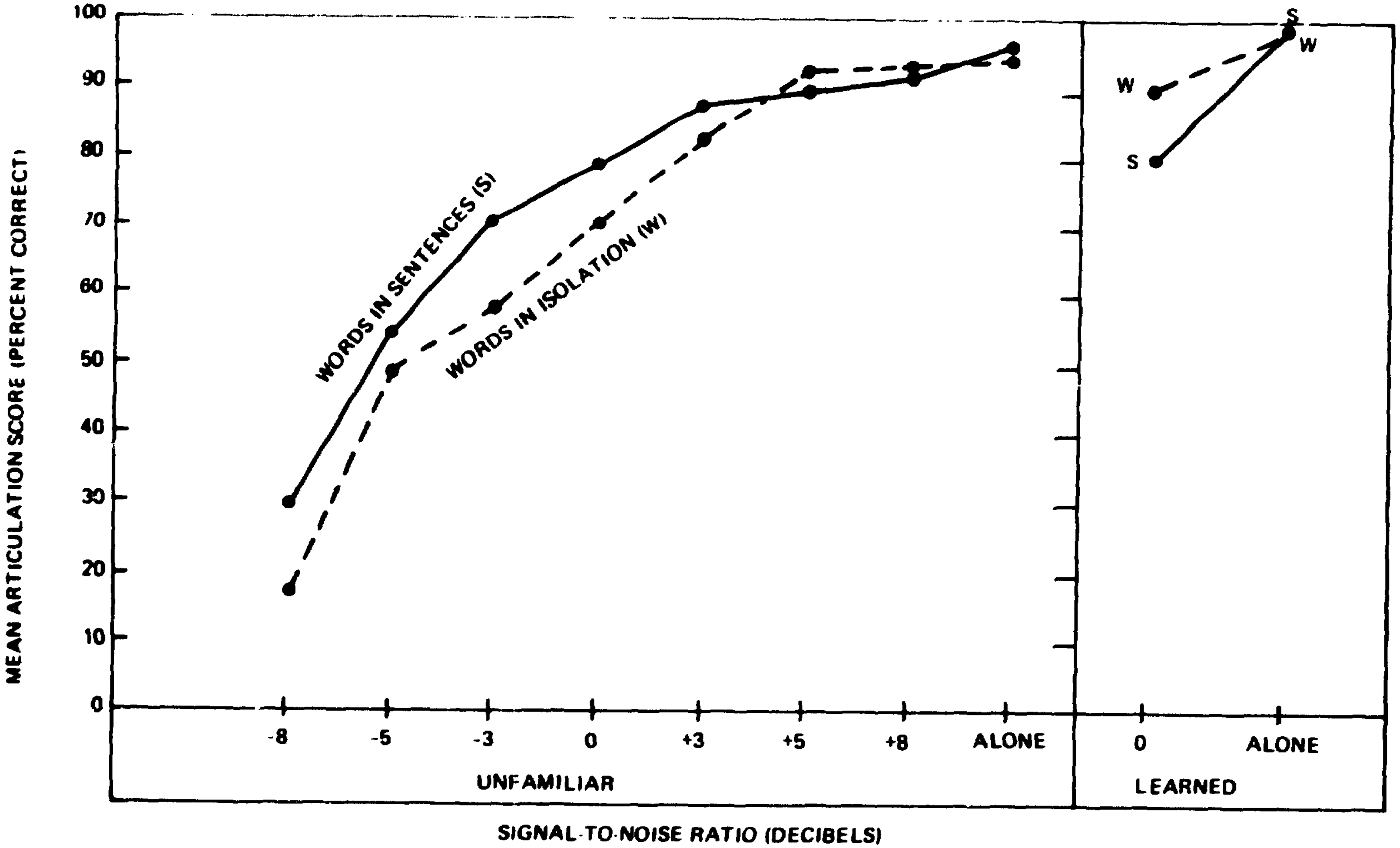
The same recommendations provided for the intensity of aural tones applies to the intensity of verbal messages (see section 3.4.1.2).

Data obtained by Boeing and Douglas indicates a strong pilot preference for automatic gain control, whereby the intensity of the aural alert or message increases or decreases as a function of the ambient noise level in the cockpit at the time of the alert.



Note Mean articulation scores for key words in synthesized speech cockpit warnings heard in background of continuous weather broadcast at seven signal-to-noise ratios and in silence. Eight airline pilots per group for unfamiliar messages, four airline pilots per group for the same messages learned before testing (Simpson, 1976)

Figure 3.4.2.6-1. Monosyllabic Words in Isolation and in Sentence Context



Note Mean articulation scores for keywords in synthesized speech cockpit warnings heard in background of continuous weather broadcast at seven signal-to-noise ratios and in silence. Eight airline pilots per group for unfamiliar messages, four airline pilots per group for the same messages learned before testing (Simpson, 1976).

Figure 3.4.2.6-2. Polysyllabic Words in Isolation and in Sentence Context

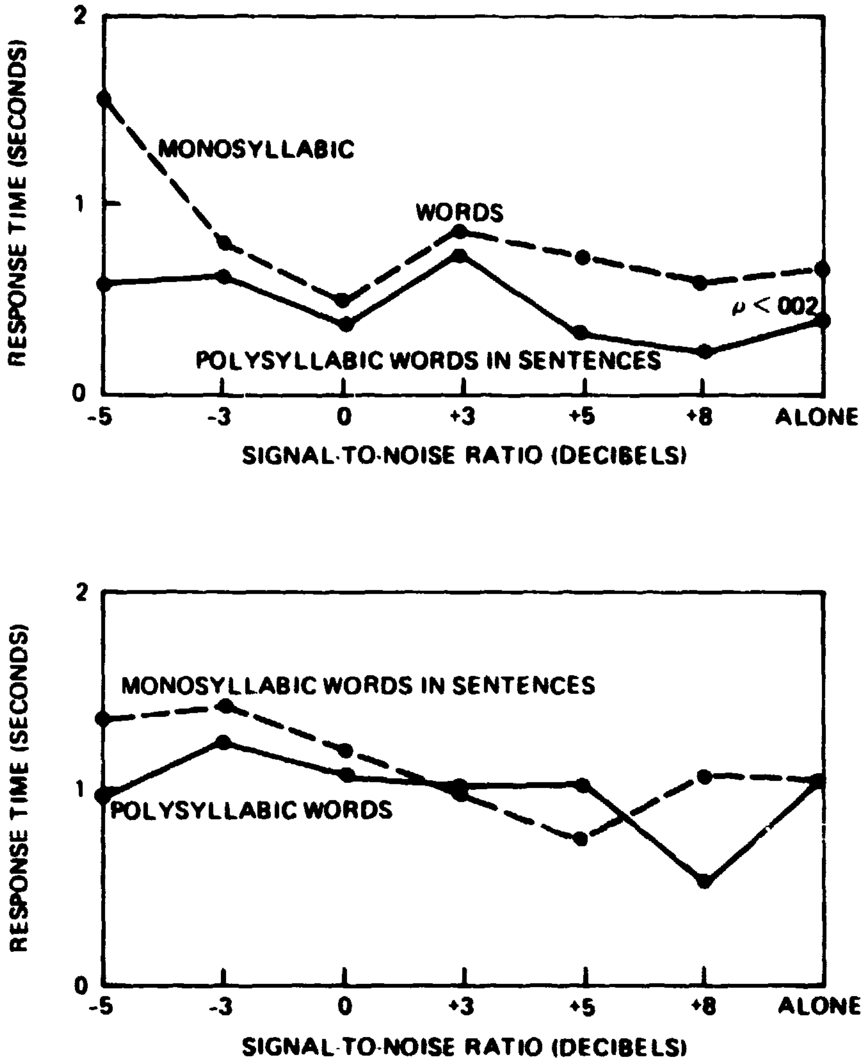


Figure 3.4.2.6-3. Response Times for Two Groups of Pilots and Messages of Different Contextual Makeup (Simpson, 1976)

3.4.2.8 MASKING EFFECTS

The masking of verbal messages should be reduced or eliminated using the same procedures described earlier in section 3.4.1.2 for auditory tones. For verbal messages masking thresholds vary as a function of voice quality; once the masking threshold has been established, minimum and optimum loudness levels should be determined for the voice model used. Figure 3.4.2.8-1 provides an example of the masking threshold values, along with the minimum, optimum and ambient noise levels as a function of voice quality used by Kerce (1979).

3.4.2.9 MODE OF PRESENTATION

The same guidelines provided for the presentation of alerting tones apply for verbal messages (section 3.4.1.3), i.e., the alerts should be presented via speakers that are approximately 90^0 from other interfering loud speakers. If headsets are worn, the verbal alerts should be presented to one earphone, and other messages to the other.

3.5 CREW OPTION AND CONTROL

3.5.1 OVERVIEW

The last function to be considered in the design of a caution and warning system is crew option and control over the system itself; the major elements of this function are:

- Inhibition
- Prioritization
- Cancellation
- Recall

Careful considerations must be made in the design of this function because by allowing the crew to inhibit, cancel, or postpone alerts there is the possibility that such control could defeat the purpose of the alerting system (Randle, Larsen and Williams, 1975). ARP 4500 contains the following:

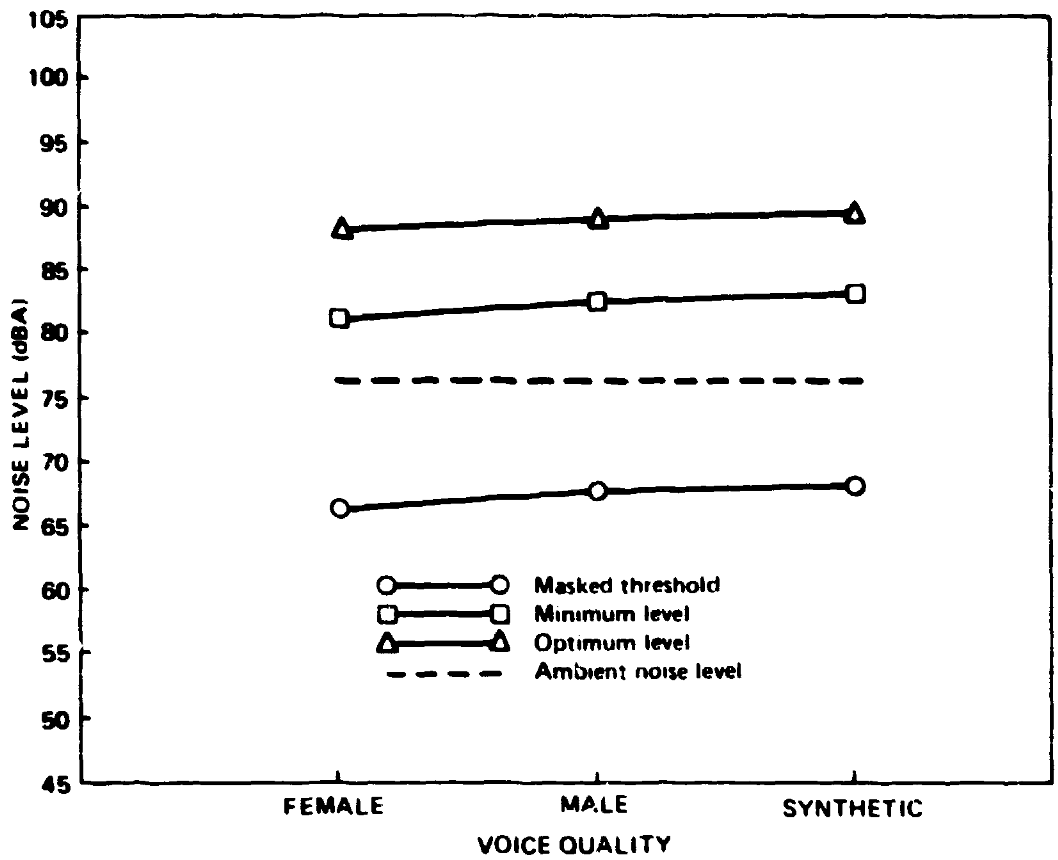


Figure 3.4.2.8-1. ***Decibel Levels for Masked Threshold and Minimum and Optimum Signal Loudness for the Female, Male, and Synthesized Voice***

- The system shall contain provisions to establish suitable priorities and to inhibit alerts under certain specific conditions.
- The system shall avoid unnecessary alerts following an intentional action such as engine shutdown.
- The design shall be such as to eliminate nuisance alerts.
- Capability to cancel visual displays of uncorrected faults and to recall such cancelled displays shall be provided.
- Alphanumeric readouts shall be self-cancelling for corrected faults.
- Capability to cancel information, advisory and caution level indications for uncorrected faults shall be provided.
- Capability shall be provided to recall readouts of uncorrected faults cancelled manually.

For the most part, the data available on crew option and control was derived from interviews with, and questionnaires administered to pilots and other flight operations personnel. This data is summarized in the following paragraphs.

3.5.2 INHIBITION AND PRIORITIZATION

Inhibition refers to delaying the onset of noncritical alerts during critical or high workload flight phases. Theoretically, prioritization should lead to a decrease in crew reation time and alert inhibition to a reduce likelihood of inappropriate action. Prioritization of alerts exists in two forms:

(1) grouping of the alerts into categories of criticality, e.g., warning, caution, advisory, etc.
(2) evaluating the importance of alerts within these categories.

3.5.2.1 PRESENT DAY UTILIZATION OF ALERT INHIBITION AND PRIORITIZATION

Alert inhibition is used on all modern commercial transport aircraft to minimize the occurrence of nuisance alerts, particularly those associated with configuration of flaps, landing gear, etc. However, very few aircraft utilize inhibits to suppress nuisance alerts from less important systems during high workload flight phases. The L-1011 and the DC-10 inhibit alerts for selected

subsystems during landing; a takeoff inhibit mode is used on the Concorde and on the A-300 to suppress all but a few critical warnings. Several late model 737's and 727's have aural alert prioritization systems.

3.5.2.2 PILOT PREFERENCE DATA

Veitengruber, Boucek, and Smith (1977) surveyed ALPA representatives, airline chief technical pilots, pilots from the Boeing, Douglas and Lockheed Flight Test organizations, and pilots from Boeing's crew training organization. The majority of these pilots felt that the potential exists for too many non-critical alerts during critical flight phases, wherein the crew cannot afford to divert their attention from the primary task of flying. The pilots were particularly concerned about distracting alerts during takeoff and landing; an inhibit scheme as shown in Figure 3.5.2.2-1 was suggested. They also suggested that inhibit schemes could be used to:

- Minimize nuisance alerts by disabling appropriate sections of the alerting system in flight phases wherein the alert has no meaning.
- Override background noise, such as radio chatter, that interferes with aural alerts.

They also felt that a method of prioritization should be developed.

Cooper (1977) conducted interviews with major airplane manufacturers and user personnel; the results of his survey agreed with the data obtained by Veitengruber, et al. Again the vast majority of the pilots surveyed stated that noncritical alerts should be inhibited during critical phases of flight. Cooper also solicited recommendations on how alerts should be inhibited. One method suggested was manual inhibiting by flight phase, relying upon crew action for implementation; another method was to use a computer with preprogrammed logic. Vanderschraaf (1976) proposed another concept called the Phase Adaptive Warning System (PAWS) wherein a switching logic module receives information from a central annunciator panel and other sensors and uses this data to inhibit and prioritize alerts.

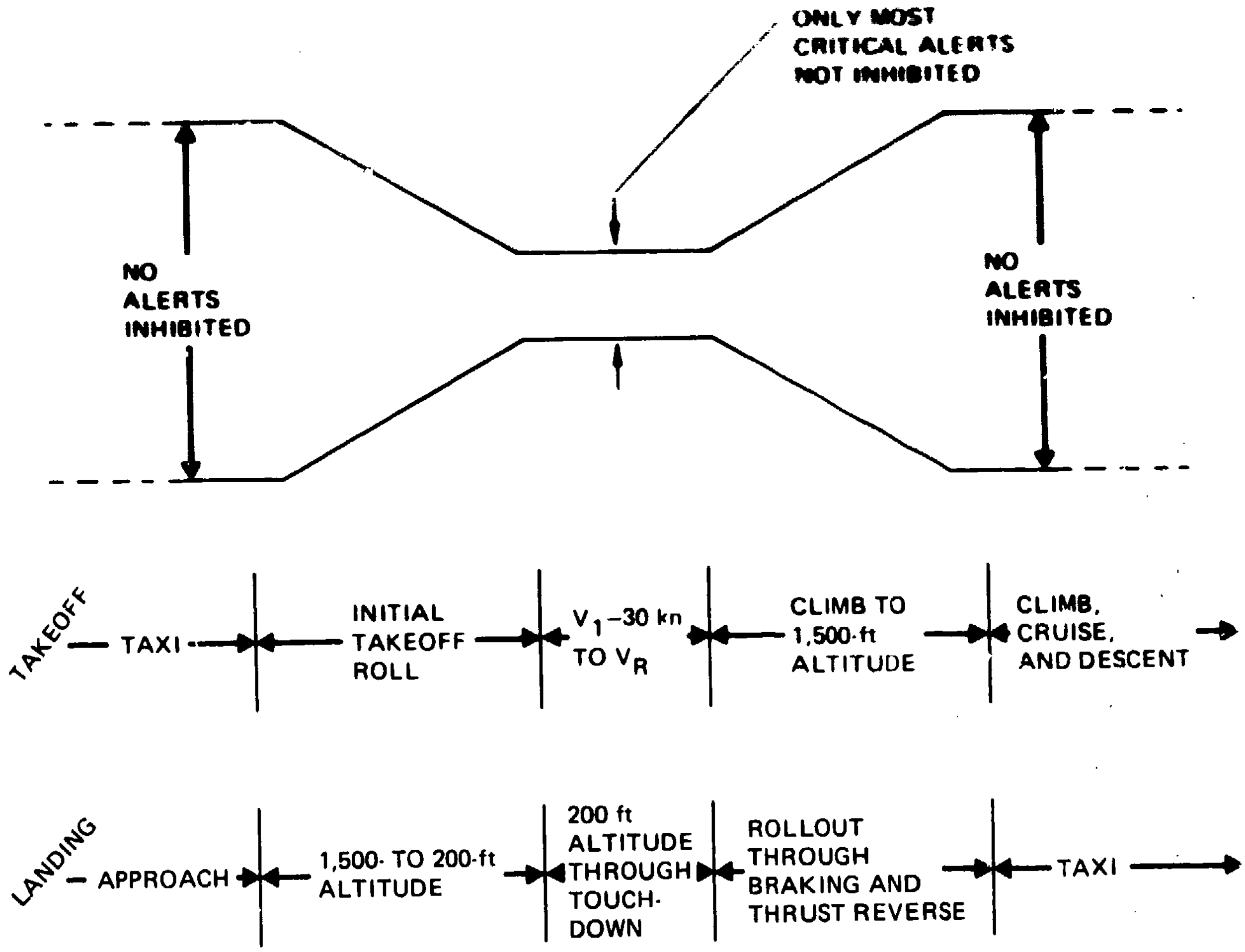


Figure 3.5.2.2-1. Alert Inhibit Scheme

The majority of the pilots surveyed felt that alert effectiveness could be improved by selective prioritization; they foresaw no serious problems as long as it was done sensibly and the pilot was informed of alerts awaiting recognition. They indicated that alerts should be grouped into three or four categories, where each category denotes a critically level; alerts within each category should also be prioritized. Also, the capability for an alert to transition from one category to another as a function of flight phase should be incorporated into the prioritization system. Although the majority of pilots favored prioritization, they could not agree upon criteria defining when prioritization was necessary (Veitengruber, et al., 1977).

Very little analytical or empirical work has been performed on how alerts can be prioritized as a function of flight phase. Veitengruber, et al, (1977) used numerical and nonnumerical methods to develop alerting categories, and to prioritize alerts within these categories as a function of flight phase. The major outputs of this work were a logic tree diagram for prioritizing alerting functions (Figure 3.5.2.2-2) and prioritization schemes for emergency, caution and advisory alerts (Tables 3.5.2.2-1 through 3.5.2.2-3). They concluded that more work is necessary to develop useful prioritization schemes, particularly in the event of multiple failures. They also concluded that since better agreement was found among pilots for high priority alerting functions, guidelines should be established only for the highest two levels, warnings and cautions, and that the prioritization of lower level alerts be left up to the airframe manufacturers and operators.

3.5.3 CANCELLATION AND RECALL

Cancellation refers to the clearing of master alerts and uncorrected fault indications. The desirability of cancelling master alerts stems from the fact that any signal which is sufficiently attention-getting to alert crew members also has the potential for creating a highly distractive environment. It is the opinion of many that this conflict can be resolved only by enabling the pilot to cancel the warning signal once it has accomplished its primary alerting function (Cooper, 1977). Cancellation of dedicated annunciators or fault indications on a central CAWS display allows the flight crew to clear the display of non-critical faults and is consistent with the philosophy of a

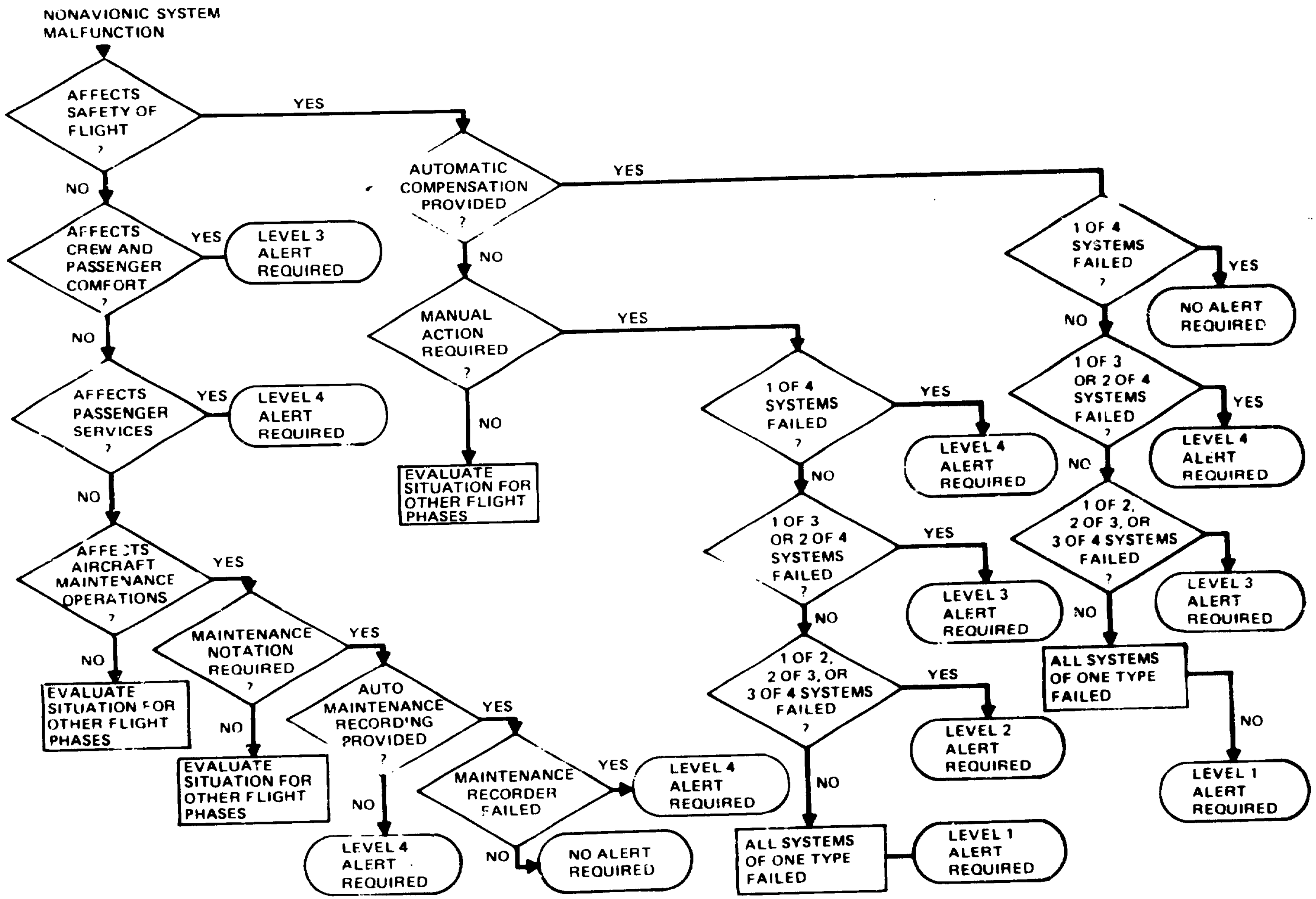


Figure 3.5.2.2-2. General Type of Logic Required to Prioritize Alerting Functions

Table 3.5.2.2-1. Example Application of Alerting Function Prioritization

Alert level (category)		1. Emergency (warning)
Alert priorities as function of flight phase	Ground maintenance	1. Gear down and locked but lever not in down detent 2. Unsafe takeoff configuration 3 Stall warning 4. Ground proximity warning
	Pre-flight	1. Gear down and locked but lever not in down detent
	Engine start	1. Gear down and locked but lever not in down detent
	Taxi	1. Gear down and locked but lever not in down detent
	Initial takeoff roll	1. Unsafe takeoff configuration 2. Gear down and locked but lever not in down detent
	Final takeoff roll	
	Initial climb	1. Stall warning 2. Ground proximity warning
	1,500- to 14,000-ft altitude	1. Stall warning 2. Ground proximity warning
	Above 14,000 ft	1. Stall warning 2. Ground proximity warning 3. Pressurization failure
	Approach (1,500- to 200-ft altitude)	1 Stall warning 2. Ground proximity warning 3. Gear down and locked but lever not in down detent 4. Unsafe landing configuration
	Landing (below 200 ft)	1. Stall warning 2. Ground proximity warning 3 Gear down and locked but lever not in down detent 4. Unsafe landing configuration 5. Autopilot disconnect
	Taxi and shutdown	1 Gear down and locked but lever not in down detent

Note Alerts prioritized as numbered, number 1 has highest priority

Table 3.5.2.2-2. Example Application of Alerting Function Prioritization (Continued)

Alert priorities as function of flight phase	Alert level (category) 2. Abnormal (cautions)						
Ground maintenance	1. Fire warning • Master • Engine • APU • Cargo	2. EGT overtemp 3. Bleed air temp high 4. Oil temp high 5. Air-conditing duct overheat	6. Passenger oxygen systems on 7. Below glide-slope warning 8. Equipment tire burst 9. Excessive airspeed or mach				
Preflight	1. Fire warning • Master • Engine • APU • Cargo	2. Air-conditing duct overheat 3. Passenger oxygen system on					
Engine start	1. Fire warning • Master • Engine • APU • Cargo	2. EGT overtemp 3. Bleed air temp high 4. Oil temp high 5. Air-conditing duct overheat					
Taxi	1. Fire warning • Master • Engine • APU • Cargo • Wheelwell	2. EGT overtemp 3. Bleed air temp high 4. Oil temp high 5. Air-conditing duct overheat	6. CSD or IDG oil temp high 7. Hydraulic pressure low • Elevator • Rudder • Aileron	8. Yaw damper fail 9. Radio altimeter fail 10. Passenger oxygen system on 11. Parking brake on			
Initial takeoff roll	1. Fire warning • Master • Engine • APU • Cargo • Wheelwell	2. EGT overtemp 3. Oil temp high	4. Hydraulic pressure low • Elevator • Rudder • Aileron	5. Stabilizer out of trim 6. Gyro fail 7. Radio altimeter fail			
Final takeoff roll							
Initial climb							
1,500 to 14,000 ft altitude	1. Fire warning • Master • Engine • APU • Cargo • Wheelwell	2. Pressurization fail 3. EGT overtemp 4. Bleed air temp high 5. Oil temp high	6. Air-conditing duct overheat 7. APU EGT overtemp 8. APU oil temp high 9. PD oil pressure low	10. APU oil quantity low 11. Hydraulic pressure low • Elevator • Rudder • Aileron	12. Yaw damper fail 13. Equipment tire burst 14. Gyro fail 15. CSD or IDG oil temp high	16. CSD or IDG oil pressure 17. Autopilot disconnect 18. Autothrottle disconnect 19. Excessive airspeed or mach	20. Gear not down and thrust lever at idle 21. Radio altimeter fail 22. Speed brake do not arm 23. Passenger oxygen on
Above 14,000 ft	1. Fire warning • Master • Engine • APU • Cargo • Wheelwell	2. EGT overtemp 3. Bleed air temp high 4. Oil temp high 5. Air-conditing duct overheat	6. APU EGT overtemp 7. APU oil temp high 8. APU oil pressure low 9. APU oil quantity low	10. Hydraulic pressure low • Elevator • Rudder • Aileron 11. Yaw damper fail	12. Equipment tire burst 13. Gyro failure 14. CSD or IDG oil temp high 15. ESD or IDG oil pressure low	16. Autopilot disconnect 17. Autothrottle disconnect 18. Excessive airspeed or mach 19. Gear not down and locked and thrust lever at idle	
Approach 1,500 to 200 ft altitude	1. Fire warning • Master • Engine • APU • Cargo • Wheelwell	2. Below glide-slope warning 3. EGT overtemp 4. Bleed air temp high 5. Oil temp high 6. Air-conditing duct overheat	7. APU EGT overtemp 8. APU oil temp high 3. APU low oil pressure 10. APU oil quantity low	11. Hydraulic pressure low • Elevator • Rudder • Aileron 12. Yaw damper fail	13. Equipment tire burst 14. Gyro failure 15. Radio altimeter fail 16. Glide-slope info fail 17. Localizer info fail	18. Autopilot disconnect 19. Autothrottle disconnect 20. Gear not down and locked and thrust lever at idle 21. Brake pressure abnormal	22. Reverser accumulator pressure low 23. Speed brake do not arm 24. CSD or IDG oil temp high 25. CSD or IDG oil pressure low
Landing below 200 ft	1. Hydraulic pressure low • Elevator • Rudder • Aileron	2. Gyro failure 3. Radio altimeter failure 4. Brake pressure abnormal 5. Reverser accumulator pressure low					
Taxi and shutdown	1. Fire warning • Master • Engine • APU • Cargo • Wheelwell	2. EGT overtemp 3. Bleed air temp high 4. Oil temp high 5. Air-conditing duct overheat	6. Brake pressure abnormal 7. CSD or IDG oil temp high 8. CSD or IDG oil pressure low				

Note: Alerts prioritized as numbered, number 1 has highest priority.

Table 3.5.2.2-3. Example Application of Alerting Function Prioritization (Concluded)

	Alert level (category)	3. Advisories	4. Information (not part of integrated warning system)
Alert priorities as function of flight phase	Ground maintenance		
	Preflight		
	Engine start		
	Taxi		
	Initial takeoff roll	Function of aircraft design Priorities to be determined by airframe manufacturer and operator	Function of aircraft design Priorities to be determined by airframe manufacturer and operator
	Final takeoff roll		
	Initial climb		
	1,500- to 14,000-ft altitude		
	Above 14,000 ft		
	Approach (1,500- to 200-ft altitude)		
	Landing (below 200 ft)		
	Taxi and shutdown		

quiet, dark cockpit, wherein no alerts are presented unless required for aircraft safety or operability. A manual recall capability is necessary to enable the flight crew to redisplay uncorrected faults.

3.5.3.1 PILOT PREFERENCE DATA

Most of the pilot preference data that is available on the subject of cancellation and recall deals with master visual and auditory alerts. In studies conducted by Williams and Simpson (1976) and Cooper (1977) the following statements were agreed to be the majority of pilots surveyed.

- Pilots can handle effectively only one emergency at a time; additional warnings can distract them from the task at hand.

- Warnings which cannot be cancelled except by fault correction may force pilots to take precipitous action which could be erroneous.

- Extremely loud or visually distractive alerting systems can interfere with cockpit communication, decision making, and crew coordination.

Mainly because of these reasons there was a general agreement that most warnings should be cancellable. However, while most pilots wanted to be able to cancel the master alert, they preferred that the fault indications remain on until the fault was corrected or the alert was cancelled by the pilot. No consensus was reached on the subject of cancelling high priority warnings. Some pilots indicated that these warnings should not be cancelled until the problem was corrected; others stated these alerts should be cancellable but that stringent criteria should be specified before implementation.

4.0 TEST FACILITY

The various study requirements dictated a facility in which flight deck system integration can be demonstrated, tested and evaluated in simulated environments. This facility consists essentially of a development cab that serves as an "operational breadboard" for the development of flight deck system concepts, functional capabilities, and cockpit interface features. Proposed systems, system changes, and alternative mechanizations can therefore be evaluated and demonstrated in this facility. This type of facility also provides a flexible experimental simulation laboratory that allows for easy introduction of new hardware and change to the flight deck system configuration. System software is modularized to facilitate change; interface equipment is flexible and thus allows for wide varieties of engineering developmental evaluations. Key elements of this type of facility have been developed in the Boeing Company's Renton Simulation Center and at Douglas Aircraft Digital Equipment Technology Analysis Center (DETAC). See Figures 4.0-1 and 4.0-2 for illustrations of these facilities. For a more detailed description refer to Appendix A.

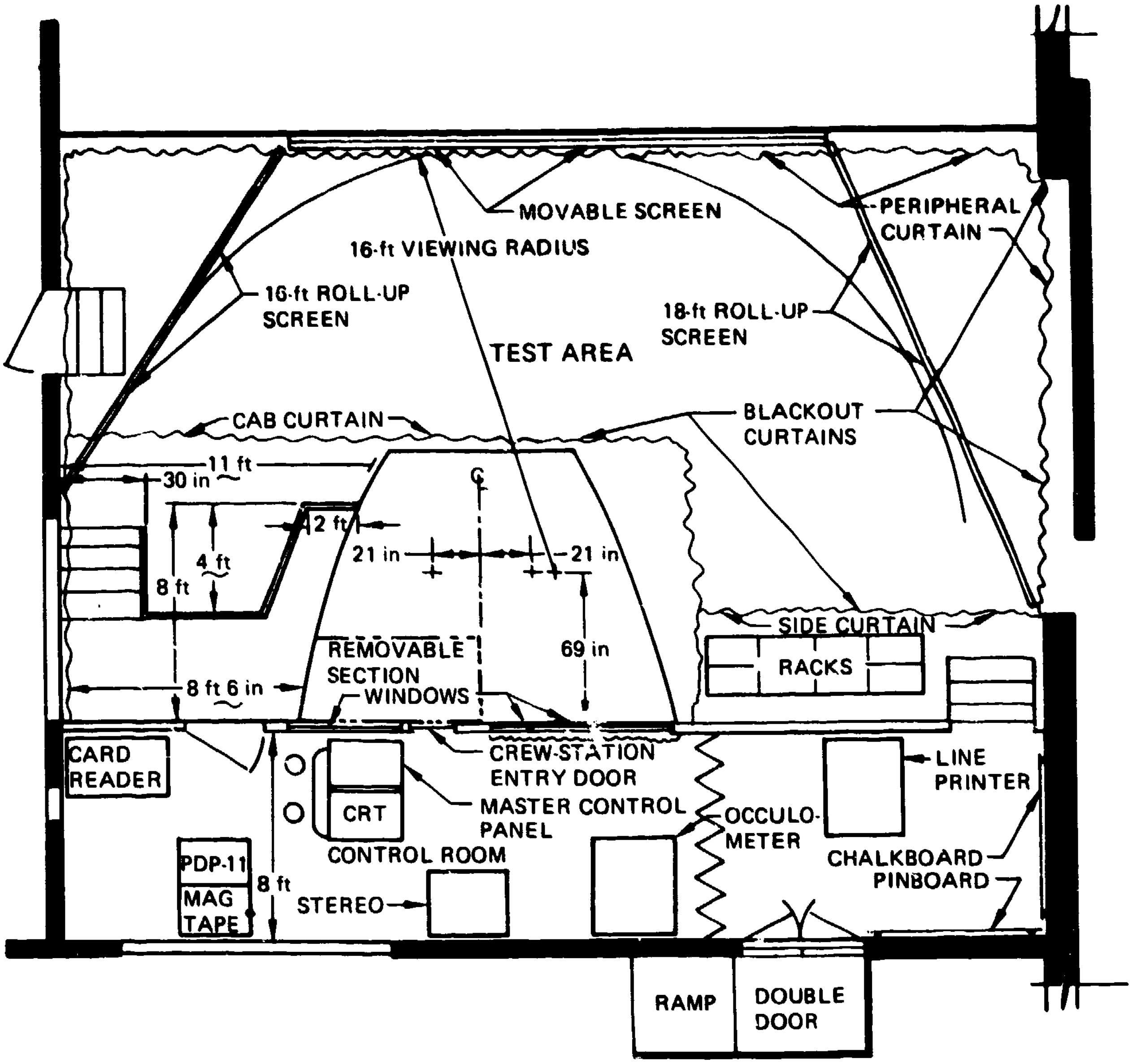


Figure 4.0-1. Boeing's Crew Systems Flight Deck Laboratory

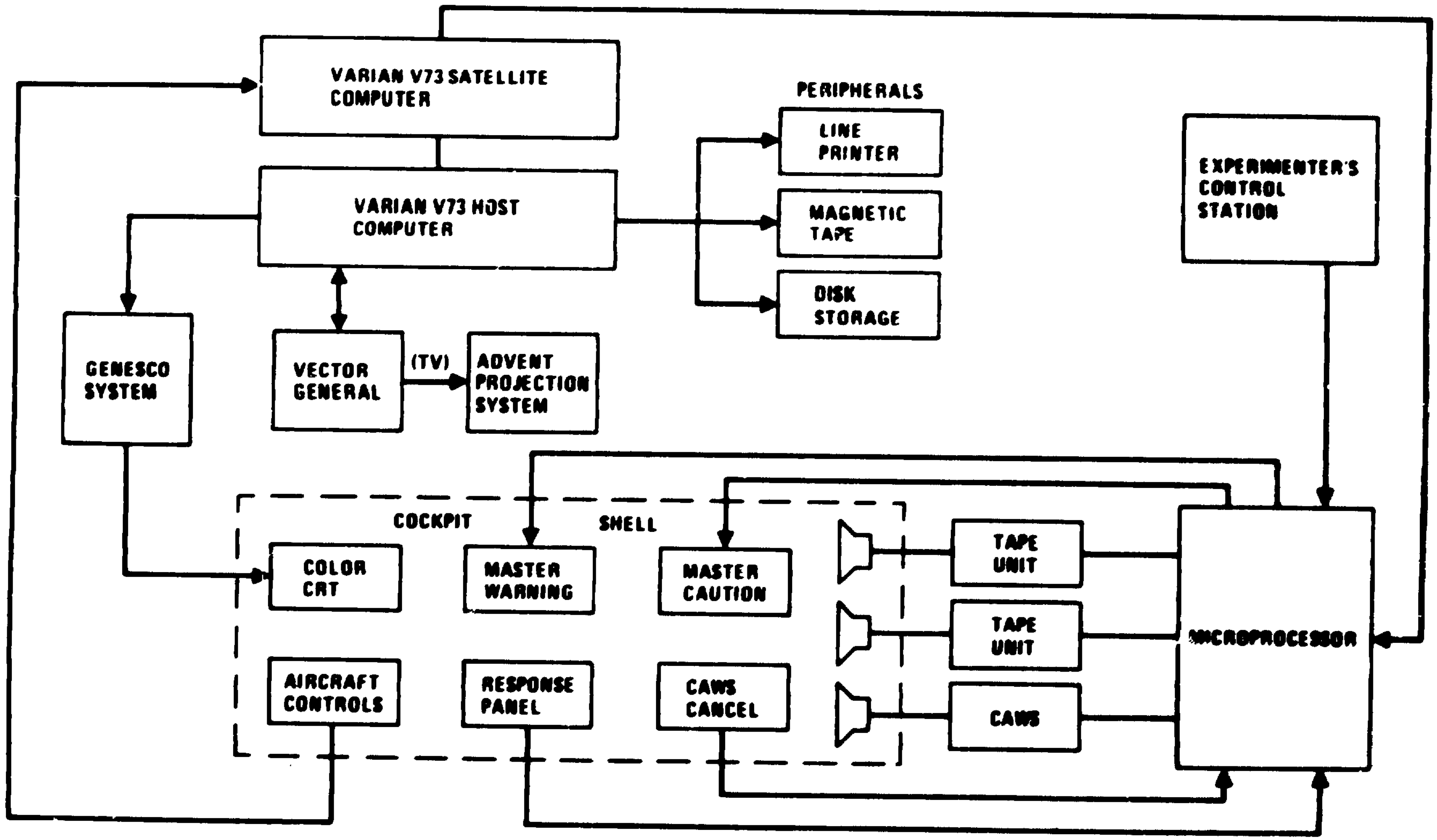


Figure 4.0-2. Douglas DETAC Hardware Configuration for Tests 3 and 4 of Aircraft Alerting Systems Standardization Study

5.0 TEST DESCRIPTION AND RESULTS

The primary purpose of the Phase I tests was to augment the existing data base on warning and caution signal detection and response. The following sections will describe in detail the tests performed and the results obtained. Five tests were conducted: tests 1 through 4 were objective tests performed in flight simulators; test 5 was designed to gather subjective data about some of the alerting parameters which did not lend themselves to simulation testing. The four objective tests investigated the two primary modes of transmitting information to the pilot, visual and auditory. The visual tests were conducted at the Boeing facility and the auditory tests at Douglas Aircraft. Unless otherwise stated, the procedures and equipment were the same for the different tests.

5.1 TEST OBJECTIVES

5.1.1 TEST 1 – VISUAL SYSTEMS

The first test, conducted at Boeing was designed to evaluate formats for the master visual attenson and the information display. The test was designed to examine the following experimental questions concerning the visual components of an alerting system:

1. Does the use of a master visual alert have an effect on detection and response time?

2. Is the attention-getting value of the master visual alert affected by the format of the alert?

3. Does the amount of information presented in the master visual alert influence response time?

4. Are the response and detection times for advisory alerts (which had no master alert) affected by the format of the master visual alert used for cautions and warnings?

5. Do any of the master visual alert formats have a disruptive effect on flight performance?

6. Does pilot workload have an effect on the detection of and response to alerts?

7. Should alert messages be grouped by urgency or should they be presented in the order which they occur?

8. Does the urgency of the alert affect response?

9. Do responses change for different combinations of master alert and information formats?

10. Is the flight task compromised by any combination of master alert and information formats?

11. Do the attenson or display formats have an effect on the number of missed alerts?

5.1.2 TEST 2 — VISUAL SYSTEMS

The second test, also conducted at Boeing was directed toward evaluating attention-getting devices and the locations for the information display.

The test was designed to examine the following experimental questions·

1. Does the use of a master visual alert have an effect on detection and response time? on missed alerts?

2. Can the absence of a master attenson be compensated for by a flashing box around the most recent alert on the information display?

3. Do pilots perform better with a combination of a master alert and a flashing box than with either one individually?

4. Do any of the attention-getting devices have a disruptive effect on flight performance?

5. Does pilot workload have an effect on detection time? on response time? on missed alerts?

6. Does the location of the information display have an effect on detection time? on response time? on missed alerts?

7. Is there an interaction between the location of the information display and the attention-getter?

8. Does urgency of the alert have an effect on pilot response?

9. Is the flight task compromised by any of the variables individually or in combination?

5.1.3 TEST 3 — AURAL ALERTS

The third test was conducted at Douglas and was directed toward evaluating the confusion potential between competing voice messages, alert message formats, the effectiveness of a precursor tone and the effect of the type of ATC voice (male/female) occurring.

With respect to optimization of the Synthetic Voice Alerting System, the following basic questions were identified:

1. Is there a significant potential for confusion and/or masking between voice alert and ATC communications sources (given conditions optimized for selective attention)?

2. Is the intelligibility of the voice alert message influenced by the quality of the competing speech source (i.e., is the voice model selected equally intelligible with respect to male and female controllers)?

3. Is the attention-getting value of a voice alert enhanced by preceding the voice message with an alerting tone (aural attenson) with regard to:

   a. time to identify and respond to failure annunciations?

   b. perception and correct interpretation of ATC communications?

4. Does the presence of an alerting tone constitute an additional distraction with respect to performance of concurrent flight crew tasks such as:

   a. safe and accurate control of the aircraft?

   b. perception and correct interpretation of ATC communications?

5. Is the effectiveness of a voice alert improved by providing redundant language in the message structure?

6. What is the optimum combination of voice alert mode and message format?

   a. voice only - word/phrase format

   b. tone-voice - word/phrase format

   c. voice only - sentence format

   d. tone-voice - sentence format

5.1.4 TEST 4 – COMBINED VISUAL AND AURAL SYSTEM

The fourth test was also conducted at Douglas and was aimed at evaluating alerting modes under differing workload and interference conditions. For the purpose of allocating functions to the auditory and visual alerting system elements, the following basic questions were identified:

1. What is the best strategy for allocating alerting functions to visual and auditory alerting modes? Is performance with each of the following modes consistent across normal and degraded flight environments?

a. Tone-visual

b. Tone-voice

c. Voice only

d. Tone-voice-visual

2. What is the effect on task performance for each of the alerting modes under the following conditions of:

   a. Visual Workload: level of difficulty of two-axis tracking task,

   b. Auditory Workload: presence/absence of concurrent ATC communications?

3. Assuming that a single alerting mode is best overall, what is the most effective method for presentating of time-critical alert information, considering:

   a. Maximum probability of timely and accurate response to failure annunciations,

   b. Minimum disruptive effects on performance of concurrent flight crew tasks that are critical to the safety factors of flight:

      (1). Accurate control of the aircraft,

      (2). Perception and correct interpretation of ATC communications?

5.1.5 TEST 5 — SYSTEM CHARACTERISTICS EVALUATION

Test 5 was administered to pilots by both Boeing and Douglas personnel. It was a subjective evaluation of alternative alerting system characteristics.

Test 5 was designed to investigate pilot preferences in the following areas:

1. How and where should a new message be presented on the information display?

2. If the system contains too many messages to be displayed at one time, how should they be handled?

3. Should a single space indent be used to distinguish advisories from cautions or should another coding scheme (such as color) be used?

4. What specific sounds are recognized by the pilots as stereotype aural alerts?

5. Are there specific characteristics of sounds which identify them as warnings, cautions, advisories?

6. If the alerting system has a master warning sound and a master caution sound, are there certain aircraft faults or conditions which would require their own specific sound?

5.2 EXPERIMENTAL DESIGN

The basic experimental design used in the objective tests (1 through 4) was a factorial design with repeated measures. The following section will present the specific design for each of the tests and the variables chosen for study.

5.2.1 TEST 1 – EXPERIMENTAL DESIGN

The basic test configuration for test 1 is shown in Figure 5.2.1-1. As can be seen, there were three independent variables: a) visual attenson format, b) central display format and c) workload. The visual attenson format variable had three levels: a single master alert for both warnings and cautions that remained illuminated until cancelled; a dual master alert (one for warnings, one for cautions) both remaining on until cancelled; and a dual master alert which flashed on and off until cancelled. The central display format had three levels: all the alerts (warnings, cautions, advisories) were presented on an alphanumeric display and were grouped by category with the most recent alert entering at the top of each group; the warnings grouped and the other two categories intermixed with the most recent message entering at the top of each group; all the messages intermixed with the most recent message entering at the top. The pilot workload variable had two levels-high and low workload. Thus, the 18 cells shown in Figure 5.2.1-1 represent a 3 x 3 x 2 factorial design with repeated measures since all pilots flew under every condition.

Figure 5.2.1-2 illustrates the relationship of the central display and the visual attenson; Figure 5.2.1-3 presents a typical format for the programmable messages. The alphanumeric display used was capable of presenting 12 lines of messages with 16 characters per line. Character size was 0.2 inch high by 0.1 inch wide; character separation was 0.08 inch word separation was 0.26 inch; and line separation was 0.05 inch. Warning messages were presented in red, cautions in amber and the advisories blue.

To evaluate the format concepts, the central display always had the same four alerts present (one warning, two cautions and one advisory) while the test flights were being performed (see Figure 5.2.1-3). The pilots were told to remember these alerts so that they would know when a new message occurred.

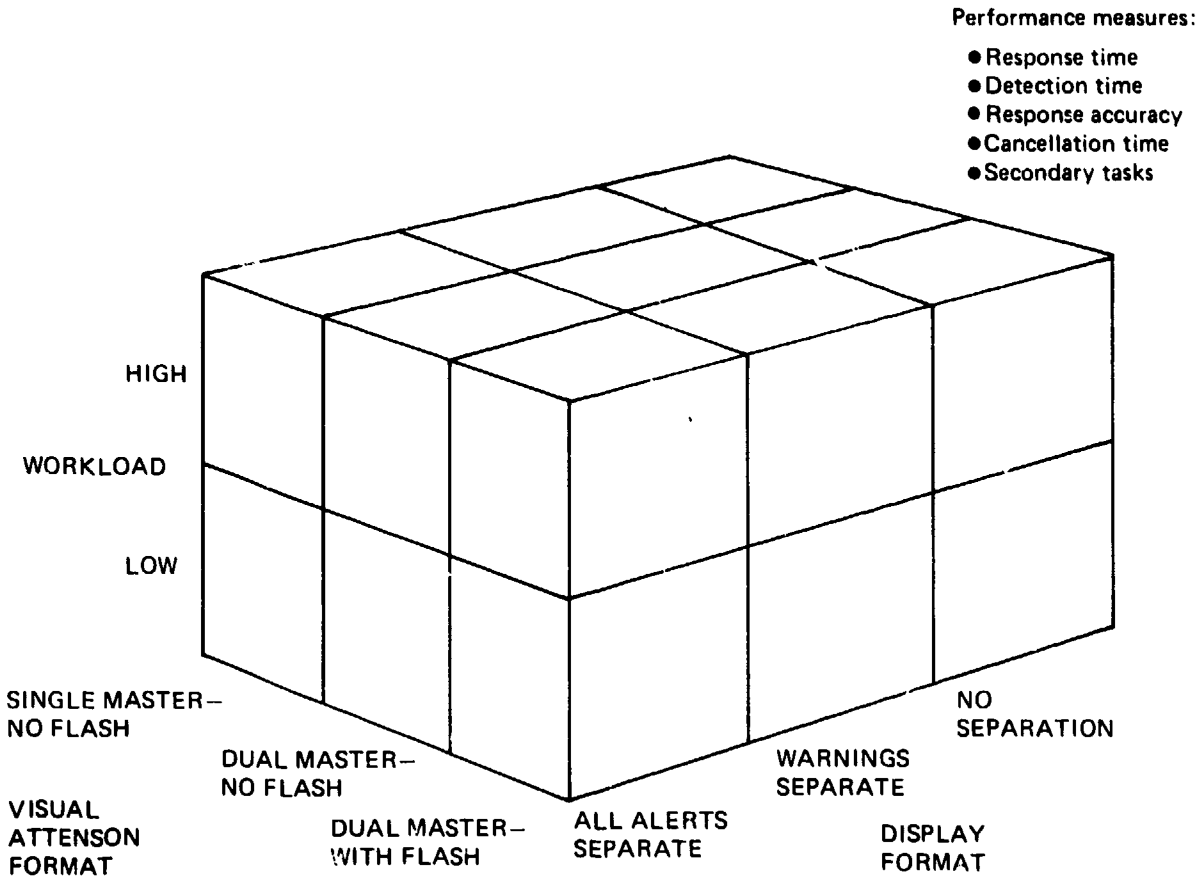


Figure 5.2.1-1. Test Configuration for Test 1

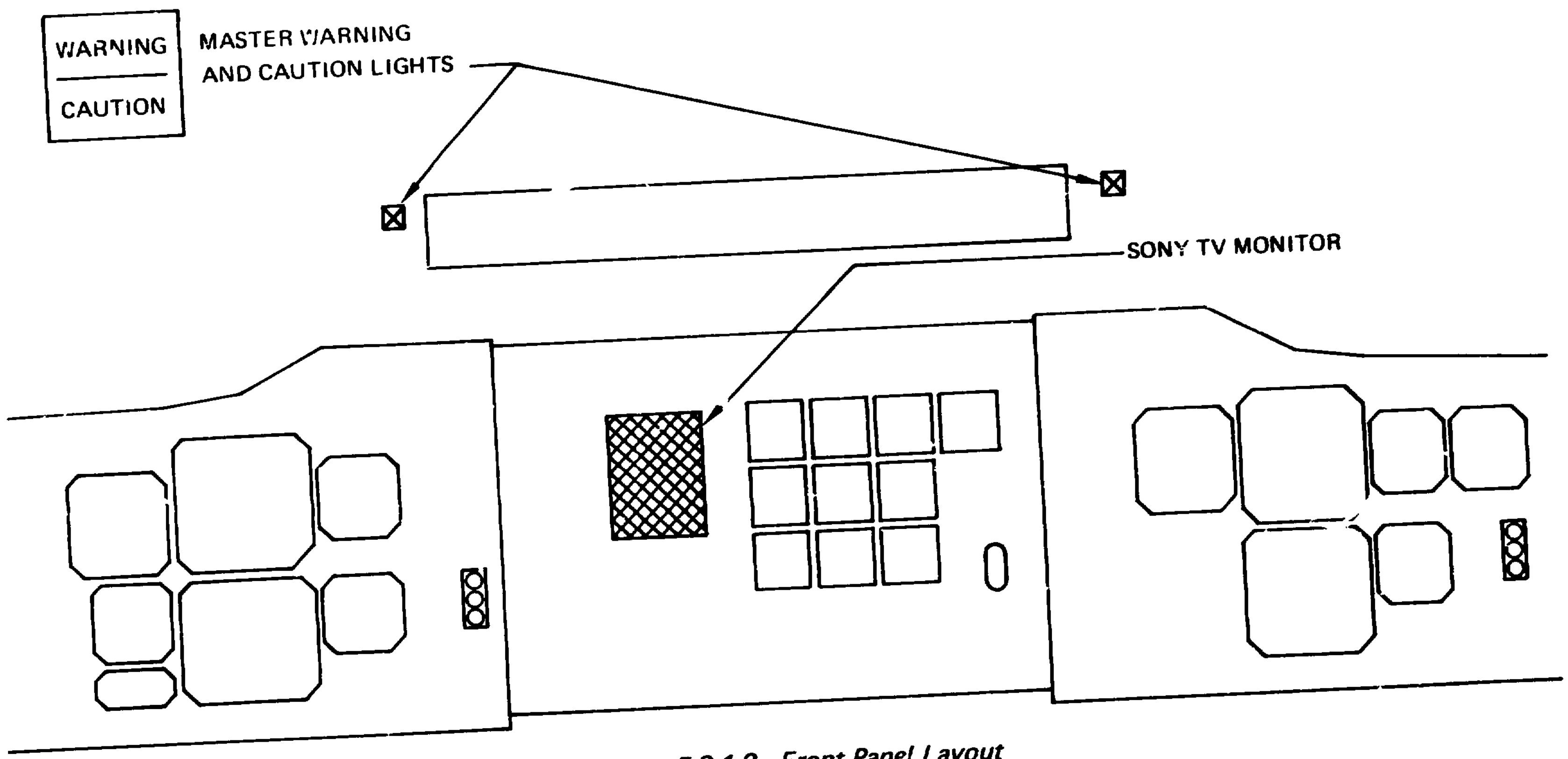


Figure 5.2.1-2. Front Panel Layout

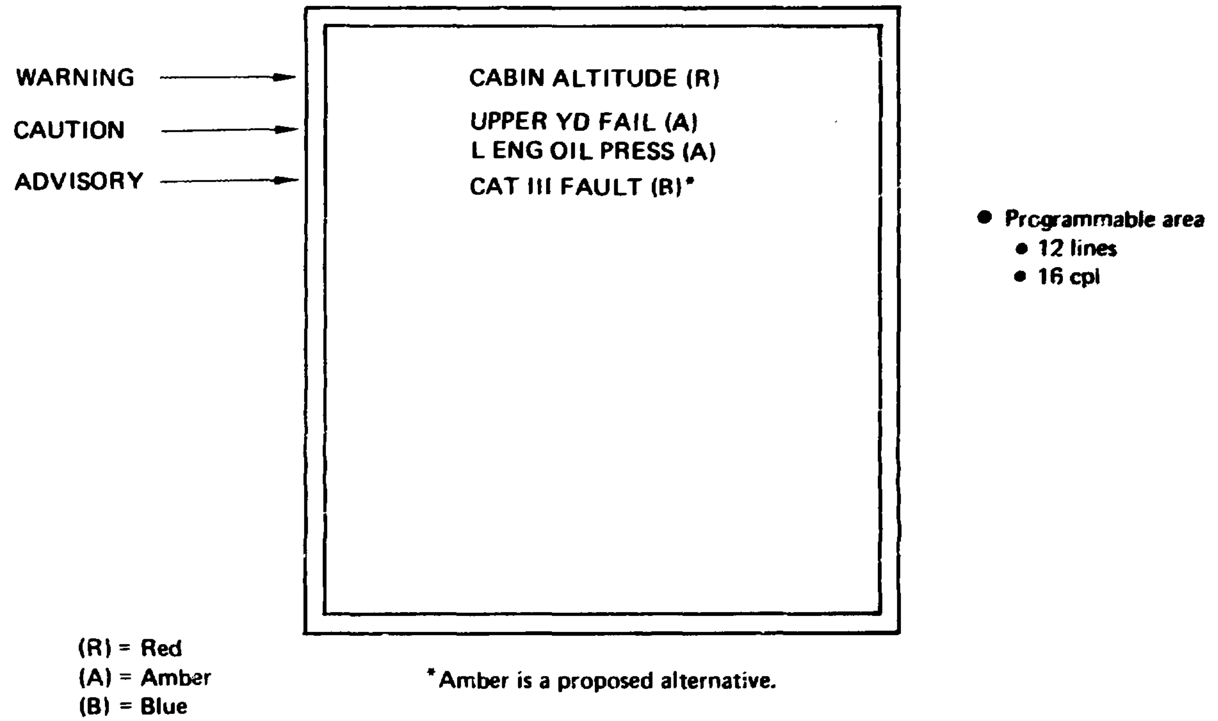


Figure 5.2.1-3. Display Format (Typical)

During the test flight the alerting system inserted the message of interest (target alert) among the existing messages according to the display format logic in use, requiring that the message either go at the top of the display or at the top of the appropriate group of messages. Since test one was concerned only with the visual system, no aural alerts were presented. This would correspond to the flight situation where the pilot is extremely overloaded in the aural channel so that the aural alert has a higher probability of being missed. The active message remained until the pilot responded to the alert at which time the message was erased. If the pilot did not respond to an alert correctly within 5 seconds of receiving the next alert, the first alert was cancelled and recorded as a missed alert; this procedure was necessary so that the number of messages on the screen would remain constant for each alert.

The master alert located directly in the pilot's primary vision area (see Figure 5.2.1-2) was divided into halves; the upper half had a red "warning" legend, the lower half an amber "caution." For the dual attenson format tests, each half could be independently illuminated depending on the alert presented. Cancellation of the master light could occur in two ways, by pressing the master alert itself or by correctly responding to the alert. For the single attenson format tests both halves of the master would illuminate for either cautions or warnings; cancellation was accomplished in the same manner as for the dual format.

For the flashing format, a rate of five times per second was used, with equal times for "on" and "off", cancellation was accomplished in the same manner as above.

The workload that the pilot encountered at the time of the alert was dependent on the flight scenario and visual scene; the high and low workload conditions were defined by the number and complexity of changes occurring during the flight as described in section 5.1.3.

The rationale for selecting the attenson and central display formats and the workload variables was as follows:

Attenson format was selected as a variable because the attenson is an extremely important component of the alerting system. The purpose of the attenson is to make the pilot aware that he has a problem; it therefore must have a high probability of being detected quickly and yet not be so intrusive that it disrupts critical ongoing tasks. It may or may not convey information about the type of problem that has occurred.

In choosing the specific attenson formats, an attempt was made to investigate the signal detection qualities, information properties and disruptive effects. The dual attenson format increases the amount of information provided to the pilot allowing him to identify the seriousness of the problem without further input. An alert must intrude to some extent on what the pilot is doing; the greater its intrusive value the higher the probability that it will be detected. This fact creates a problem in that the designer must make the trade between the probability of detection and the disruptive effect of the signal. Flashing lights have a higher intrusive quality than steady lights but the improvement in detection performance may not justify the irritation and disruption caused.

The second variable to be investigated was the central display format; the display formats were chosen to represent a range of computer sophistication and logic. In the case of separate alert groups, the logic must keep track of each group and each alert, predetermined classification and priorities must be followed. If, however, the alerts are mixed, the logic is greatly simplified in that the most recent alert as always presented at the top of the display.

Different formats pose different problems for the pilot. When alert messages are categorized by severity the pilot has color and location cues to help him find and identify an alert; however, he must continuously look at all category locations to detect the most recent message; the possibility of missing a new alert is therefore increased. If the messages are displayed in order of their occurrence, identifying new alerts becomes easier because the newest alert is always at the top; however, it becomes more difficult for the pilot to evaluate the condition of his aircraft because he must read all the messages and prioritize them himself.

The workload variable (described in section 5.1.3) was included in the testing to provide a realistic framework from which to evaluate performance.

5.2.2 TEST 2 – EXPERIMENTAL DESIGN

The basic configuration for this experiment is shown in Figure 5.2.2-1. There were four independent variables: a) visual attenson, b) central display location, c) central display cueing and d) pilot workload. The visual attenson variable had two levels; master alert present and master alert absent. The central display location variable had three levels: a single central display for all alerts located on the front panel between the Captain and First Officer (see Figure 5.2.2-2); a display between the Pilot and First Officer for cautions and advisories and warning displays in front of the pilot; finally the same configuration but the display on the middle panel is only for advisories while the other displays are for warnings and cautions. The cueing variable had two levels; the most recent message with a flashing box around it or the most recent message with no box around it. The pilot workload variable had two levels: high workload and low workload. Thus, the 24 cells shown in Figure 5.2.2-1 represents a 3 x 2 x 2 x 2 factorial design with repeated measures, with all pilots flying under every condition.

The visual displays were capable of presenting 12 lines of messages with 16 characters per line. The characters are .2 inches high and .1 inches wide. They are separated horizontally by .08 inches and each word has a horizontal separation of .26 inches. The vertical separation between lines is .05 inches. Warning messages were red in color while caution was amber and advisories were blue.

To fully evaluate the format concepts the central display always had the same four alerts present (one warning, two cautions and one advisory) while the test flights were being performed (see Figure 5.2.1-3). The pilots were told just to note the alerts so that they would know when a new message occurred. During the test flight the alerting system inserted the message of interest (target alert) among the existing messages according to the display format logic in use, requiring that the message either go at the top of the display or at the top of the appropriate group of messages. Since test one was concerned only with the visual system, no aural alerts were presented with any of the visual alerts. This would correspond to the flight situation where the pilot is extremely overloaded in the aural channel so that the aural alert has

Performance measures:
- Detection time
- Response time
- Response accuracy
- Secondary tasks

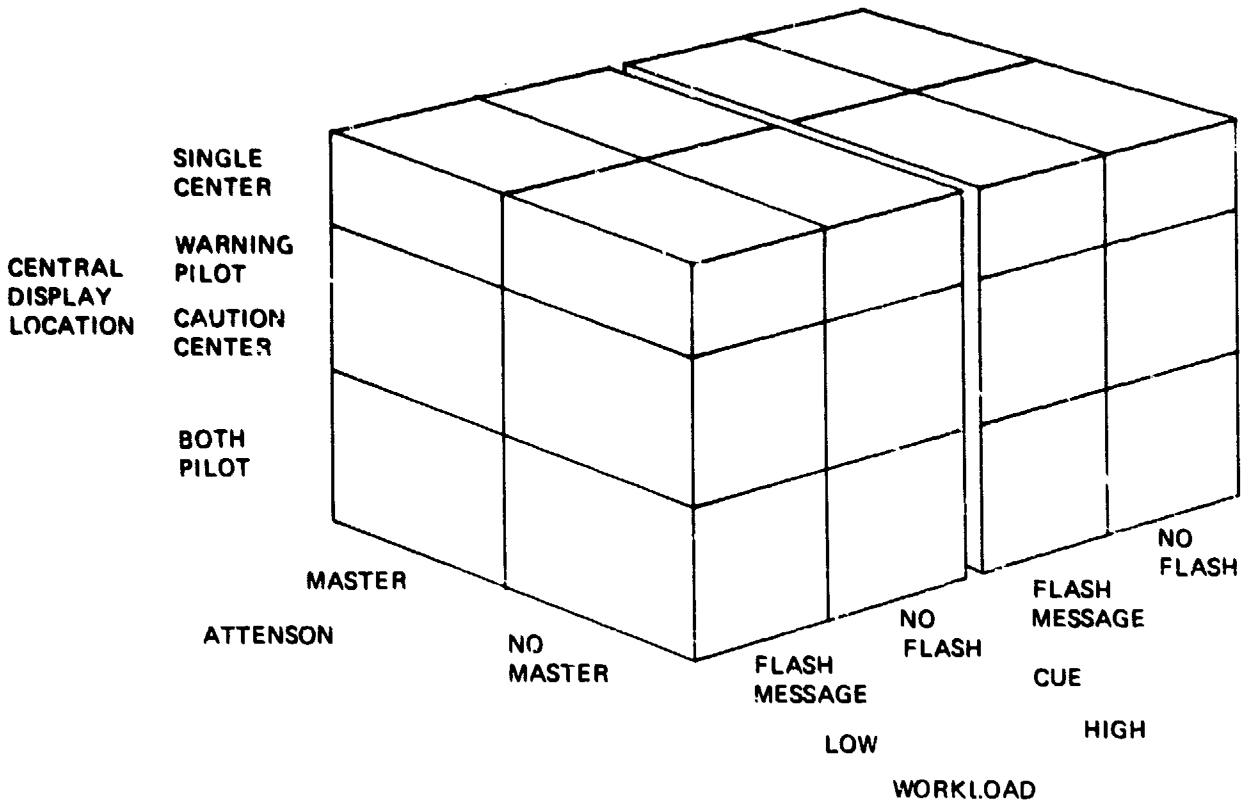


Figure 5.2.2-1. Test Configuration for Test 2

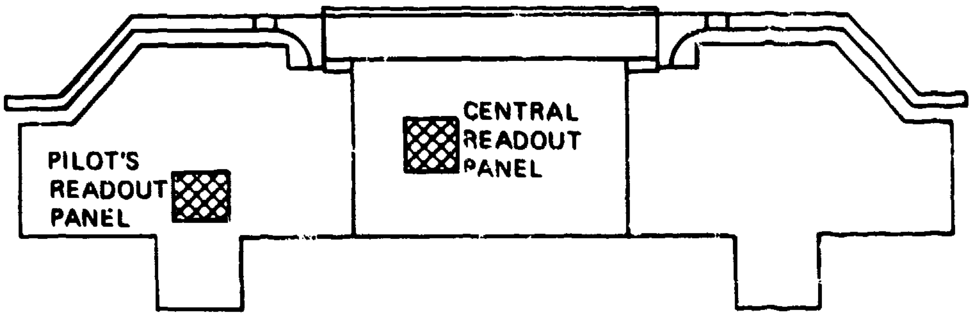


Figure 5.2.2-2. Separate Warning and Caution Displays

a higher probability of being missed. The active message remained until the pilot responded to the alert at which time the message was erased. If the pilot did not respond to an alert correctly within 5 seconds of receiving the next alert, the first alert was cancelled and recorded as a missed alert; this procedure was necessary so that the number of messages on the screen would remain constant for each alert.

This study was conducted with visual systems only; no aural alerts were presented. The active message remained until the pilot corrected the fault (responded in a prescribed manner) at which time the message was erased. If the pilot did not respond correctly within .5 seconds of receiving the next message the active message was cancelled and recorded as a missed alert. This procedure was necessary to insure that the amount of clutter (number of messages) on the screen remained constant for each alert.

The master alert was configured as seen in Figure 5.2.1-2 and cancellation of the alert occurred either when the problem was solved or when the pilot pushed the alert switch. Test 2 used the dual master alert.

When cueing was used with the information display, the box around the alert message flashed with equal on and off times.

As in test one, pilot workload at the time of the alert was dependent on the flight scenario and the outside visual scene. The high and low workload conditions were defined by the number and complexity of changes occurring during the flight (see section 5.2.3).

This experiment was designed to investigate the interrelationship between alert detection and identification; the rationale for selecting test variables was as follows:

As in the first experiment, the visual attenson was used primarily to inform the pilot that a problem existed; the attenson must be intrusive without being disruptive to critical ongoing tasks. This study was aimed at determining whether this configuration would improve detection and response performance compared to a configuration which has no master alert.

The second variable chosen for investigation was the location of the information display. It may be possible to use the alert message itself as the attenson if the message could be located in the pilot's primary field of vision; if the message itself can provide enough intrusion to have a high probability of being detected then it may be possible to eliminate the master alert. Because the message itself may not have enough impact to penetrate the pilot's workload it was decided to also investigate a cueing factor; by putting a flashing box around the most recent message it was hoped that the attention getting properties of the alert message would be increased.

Again, the workload variable was included in the testing to provide a realistic framework from which to evaluate performance.

5.2.3 TEST 3 — EXPERIMENTAL DESIGN

There seems to be general agreement that conventional auditory alerting systems place excessive demands on the information processing and memory capabilities of the crew. These systems employ as many as 12 to 17 dedicated alerting tones to identify various emergency or abnormal conditions. In an industry survey, Cooper (1977) found that the general feeling among aviation industry representatives is that flight crews of current transport aircraft are often overwarned and overworked.

Synthetic voice warning devices offer considerable promise as a replacement for or supplement to discrete tone systems. Before extensive applications of voice warning systems can be made, a number of operational issues need to be addressed. There has been some concern expressed about the expanded use of voice warnings in the cockpit because of the possibility of interference with crew or ATC communications. A number of additional design issues were identified by Simpson and Williams (1975). They include the actual message content, linguistic redundancy, signal to noise ratio, interference with concurrent tasks, and listener expectations generated by the pragmatic or real world context in which messages are presented.

The purpose of test 3 was to assess the impact of several alternative design features on the attention-getting value of voice alerts and to resolve some basic issues related to synthetic speech effectiveness. The experiment was designed to evaluate the effects of two characteristics of the alerting system (alerting tone and voice message format) and two basic features of the auditory environment (level of auditory workload and quality of the competing speech source).

Some existing guidelines and standards specify that all voice alert messages should be preceded by an attention-getting tone. Mil-Std-1472B (1974) suggests that an alerting signal or tone should precede the actual alert message. This tone should also designate the general severity of the problem (i.e., differentiate between warning, caution and advisory level alerts). It is also stated in this report that in time critical situations, the alerting tone may be omitted to expedite the fault correction process. ARP -450D

(1979) specifies that an attention-getting tone shall precede warning, caution and advisory level alerts, clearly differentiating the alert priority levels. Some evidence exists that shows superior performance to be associated with the absence of an alerting tone (Simpson and Williams, 1978). The performance criteria used in this study consisted of response times to a number of alert messages presented in the context of a simulated cockpit environment. In view of this apparent conflict in the existing literature, it was determined that additional test data is required to establish the attention-getting value of precursor tones.

A second issue which was to be resolved concerned the formatting of the voice message. The conventional approach to wording of voice alert messages is to use brief, concise words or phrases stating the nature and location of the problem. This is the methodology prescribed in document ARP-450D and Mil-Std-1472B. Other sources suggest that redundant linguistic cues provided by complete sentence messages will elicit superior performance in terms of alert recognition and response time (Boucek, Veitengruber and Smith, 1977; Simpson, 1976). This view is supported by experimental evidence generated by Simpson and Williams (1978). They found no significant difference in response time between word and sentence formatted messages when four pilots flew simulated instrument approaches. Hart and Simpson (1976) collected data showing that complete sentence messages were more intelligible than short phrase or word-formatted messages when presented against a background of simulated communications. Other studies have illustrated that response time is shorter and intelligibility higher when additional words are used in voice warnings to provide context cues (Simpson, 1976; Simpson and Hart, 1977).

A third issue addressed by test 3 concerned auditory interference. It was anticipated that the overall effectiveness of the voice warning system would be dependent on the level of auditory interference, particularly when the auditory task involves listening to conflicting speech communications. Although previous studies have dealt with intelligibility of voice alert messages, little data has been acquired regarding mutual interference and masking between voice alert and speech communications.

The last issue addressed was the quality of the competing speech source and its potential effects on performance. In order to minimize confusion and masking effects on other communications, a distinctive female speaker was selected as the voice model for the Central Aural Warning System. Since female voices are becoming more common in the cockpit environment additional data was acquired to determine if the quality of the competing speech source had a significant effect on intelligibility of the voice warning message.

The block diagram in Figure 5.2.3-1 illustrates the variables that were addressed and the performance measures used. Half of the trials called for word-formatted alert messages while the other half involved sentence-formatted messages. Manipulation of this variable provided information on the possible performance benefits gained by providing redundant language in the message structure. A second variable manipulated was ATC message onset. This variable was chosen in an effort to determine the existence of masking effects and/or confusion between voice alerts and ATC communication sources. The third variable manipulated was the ATC voice (male/female). It was necessary to determine whether the voice model selected for use in the aural warning system was equally intelligible with respect to male and female controller voices. Specifically, the female alerting voice used in the aural alerting system was introduced concurrently with either male or female ATC messages during selected trials to make this determination. Finally, the presence/absence of an alerting tone prior to the voice alert message was manipulated. While it has been suggested that a tone prior to the alert message may increase the attention-getting value of the alert, the possibility exists that the tone will cause additional distraction with respect to performance of concurrent flight crew tasks.

A 2 x 2 x 2 x 2 repeated measures design was employed, allowing each pilot to perform under all 16 conditions. This approach was taken to minimize the number of observations needed to establish the significance level of the test variables previously identified.

Voice alert messages were presented in two basic formats. The word-phrase format consisted of a brief declarative statement identifying the nature and locations of the fault. The complete sentence format provided the same

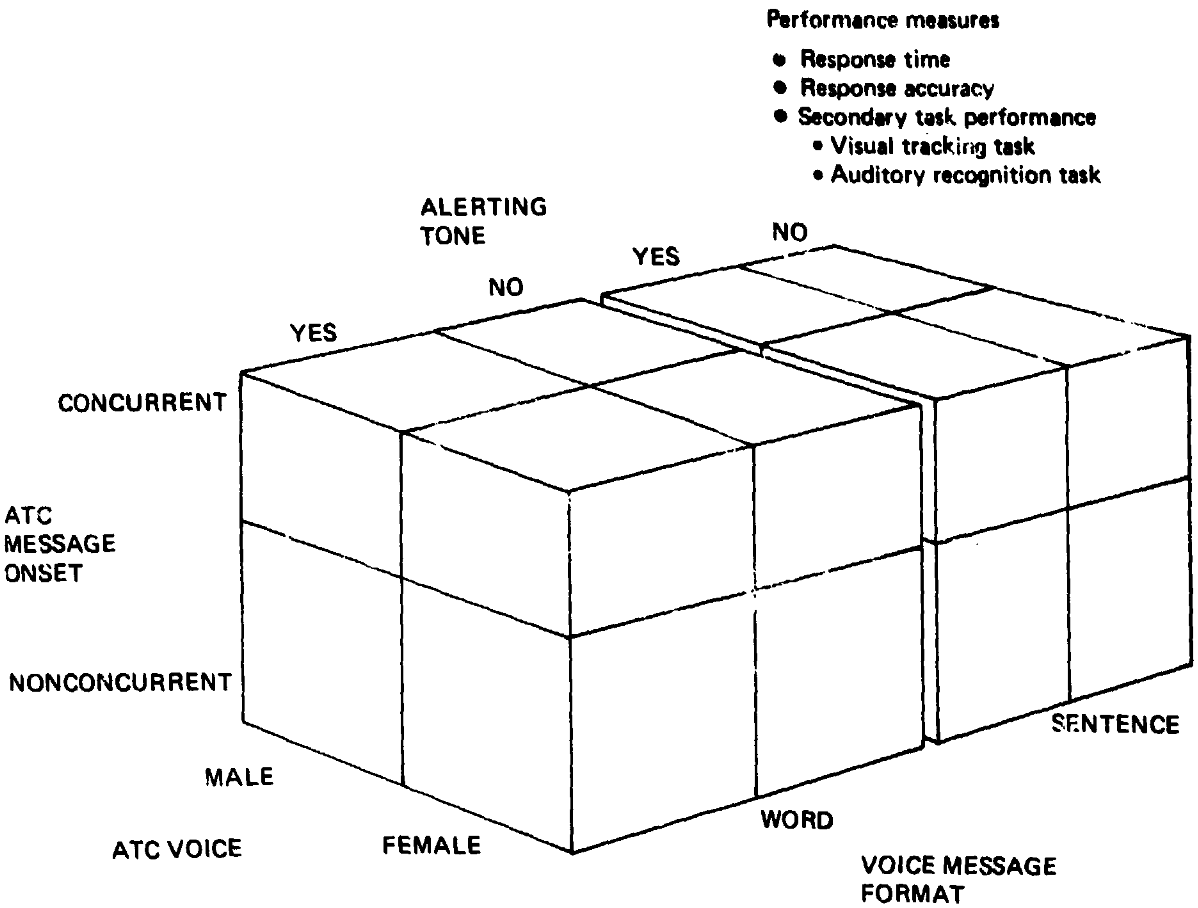


Figure 5.2.3-1. Test 3: Auditory Alerts

essential information with linguistic redundancy in the message structure. With the complete sentence format, the critical portion of the message was preceded by one or more low information content words. The alert message set and format alternatives are summarized in Appendix B.

For half of the experimental trials, aural alert messages were presented by means of voice message only. For the remaining test trials, voice messages were preceded by one of two alerting tones. The nature of the alerting tone indicated the alert priority level. Warning messages were preceded by a 750 msec high pitched horn and caution messages were preceded by a C-chord of the same duration. Alert messages were cycled in a tone-voice-tone sequence with an off-time interval of 500 msec between repetitions. All auditory alerts were annunciated continuously until cancelled by activation of the correct overhead switch. The peak loudness levels for all tone and voice alerts were adjusted to approximately 85 dBA in order to maintain the same signal-to-noise ratio established for ATC communications.

5.2.4 TEST 4 — EXPERIMENTAL DESIGN

It has been suggested that the excessive number of discrete visual and aural alerts in today's cockpit causes an excessive demand on the information processing and memory capabilities of the flight crew. There is general agreement that an improved cockpit alerting system would minimize the time required to detect, evaluate and correct impending failures. This system would also minimize the distracting effects on other flight crew tasks (i.e., aircraft control and communications).

The use of synthetic voice warning systems and programmable visual displays represent a potential solution to this problem. The advanced technology associated with these devices allows for potential operational advantages such as priority logic and flight phase adaptation. Before these systems can be incorporated into commercial aircraft, several operational and developmental issues must be resolved.

There has been some concern about the expanded use of voice warnings in the cockpit because of the possibility of interference with crew or ATC communications. Application of visual displays in the cockpit should also be approached with caution because visual media alone may be insufficient in its attention-getting ability during periods of high visual workload. Data collected by Bate (1969) showed median response times to visual warning signals to be significantly higher than those for tone-voice warning signals.

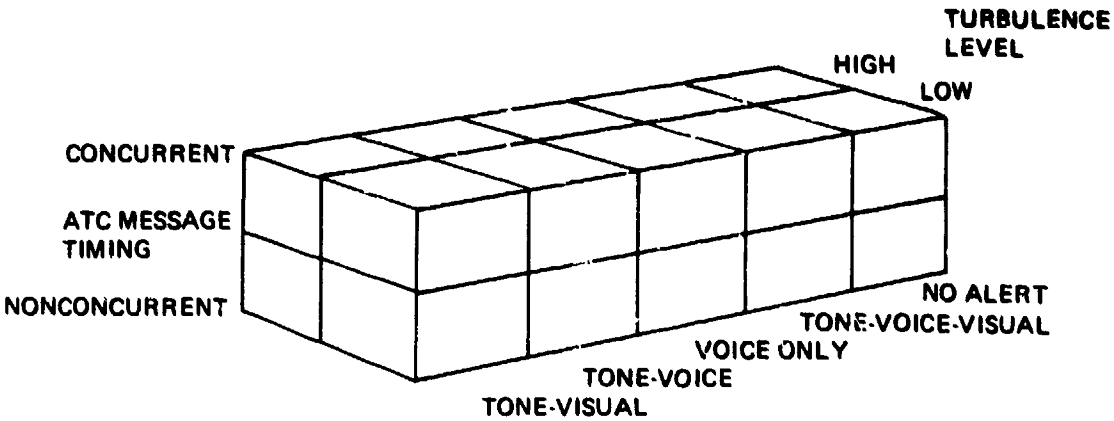
The purpose of test 4 was to investigate several possible alerting modes to identify those which best facilitate optimal flight crew performance. The potential benefits to be derived from a successful effort in this area include reduced cockpit workload as well as improved aircraft safety.

At present, there is no single alerting mode that has been given overwhelming support by the aviation community. There is a general consensus that bimodal presentation of alerting information is equal to or better than single modal presentation (Adams and Chambers, 1962; Klemmer, 1958; Bate, 1969; Mil Std 1472B, 1974). There are still a variety of viewpoints relative to which alerting modes should be employed. As mentioned earlier, Bate (1969) found

response times to be shorter with a tone-visual alerting system when compared with a visual system. Mil Std 1472B calls for voice warnings to be used in conjunction with visual displays, with the voice warnings being supportive of the visual display unit. In a survey of the commercial aviation industry, Cooper (1977) found that, while most of those questioned found voice warnings to be desirable, they also saw a need for a central visual display in the cockpit. The general feeling was that with the technological progress being made with visual displays the benefits to be derived from their use would represent a positive addition to the cockpit. There is also evidence which suggests that voice warnings, employed without an alerting tone, produce shorter response times than when an alerting tone is used (Simpson and Williams, 1978).

An additional issue to be addressed involves the effects of concurrent task loading on performance using various alerting media. It is suggested that high auditory workload would degrade alerting task performance when voice warnings are used exclusive of any supporting visual display or alerting tone. Similarly, high visual workload may be cause for degraded performance when a visual display is the only alerting media. With this in mind, it may be that a pilot selective system may be desirable where the alerting media to be selected would be dependent on the auditory and visual workload.

The experimental variables are shown in Figure 5.2.4-1 along with the evaluation criteria. As can be seen, half of the trials were performed under conditions at low turbulence while the remainder were conducted under high turbulence conditions; this variable was manipulated to determine if turbulence level would affect alert task performance as well as flight task activity. A second variable manipulated was the ATC message onset. This variable was chosen to expose any masking effects and/or confusion between voice alerts and ATC communication sources. The third and primary variable used in test 4 was the alerting mode. It was necessary to determine which combination of alerting media would be most conducive to optimal pilot performance. As mentioned earlier, it was thought that performance would vary for each alerting mode as a function of the level of visual and auditory task loading. A mode containing no alert messages was added to the design to provide a baseline configuration against which the experimental alerting modes could be compared.



- Evaluation criteria
    - Response time
    - ATC readback accuracy
    - Tracking task performance
    - Pilot preferences

Figure 5.2.4-1. Test 4: Comparative Evaluation of Alerting Modes

A 2 x 2 x 5 repeated measures design was employed, allowing each pilot to perform under all 20 conditions illustrated in the block diagram shown in Figure 5.4.1. This approach was taken to minimize the number of observations needed to establish the significance levels of the variables.

Fault messages were presented by means of four different alerting modes:

Tone - Visual

Tone - Voice

Voice Only

Tone - Voice - Visual

With the exception of the voice only alerting mode, messages were preceded by one of two alerting tones. The nature of the alerting tone indicated the alert priority level. Warning messages were preceded by a 750 msec high pitched horn and caution messages were preceded by a C-chord of the same duration. Aural alert messages were cycled continuously with an off-time interval of 500 msec between repetitions. All auditory and visual alerts were annunciated until cancelled by activation of the correct overhead switch. The peak loudness levels for all tone and voice alerts were adjusted to approximately 85 dBA to maintain the same signal-to-noise ratio established for ATC communications.

5.2.5 TEST 5 – SUBJECTIVE EVALUATION DESIGN

5.2.5.1 VISUAL DISPLAY EVALUATION

Although identical evaluations were performed by Boeing and Douglas, differences in the data collection procedures made it necessary to report the data separately. These differences will be noted in the following sections; where differences are not discussed, it may be assumed that the data collection procedures were identical.

Three conceptual areas relating to the information display were evaluated: display format, overflow logic and advisory color.

5.2.5.1.1 DISPLAY FORMAT

Three formats were considered. The first displayed the alerts by category, the most recent alert presented at the top of its category (see Figure 5.2.5.1.1-1). The warning category appeared at the top of the display, cautions in the middle and advisories on the bottom. The order of occurrence of alerts could therefore be determined only within categories. An alternative to this scheme had no categorization; the most recent alert always appeared at the top of the screen and shifted the existing messages down one line (Figure 5.2.5.1.1-2). Another alternative presented the warning category at the top of the display; cautions and advisories grouped at the bottom as they occurred.

5.2.5.1.2 OVERFLOW

Three concepts for accommodating more messages than could be displayed at one time were considered. The first was the straight overflow concept wherein the oldest message dropped off the bottom of the screen. The only way the pilot could recall a lost message was to correct a more recent fault; then the last alert to leave the screen would be the first to return.

The second concept provided the pilot a means to scroll the messages up or down on the display for access to all the alerts; message lines (cautions and

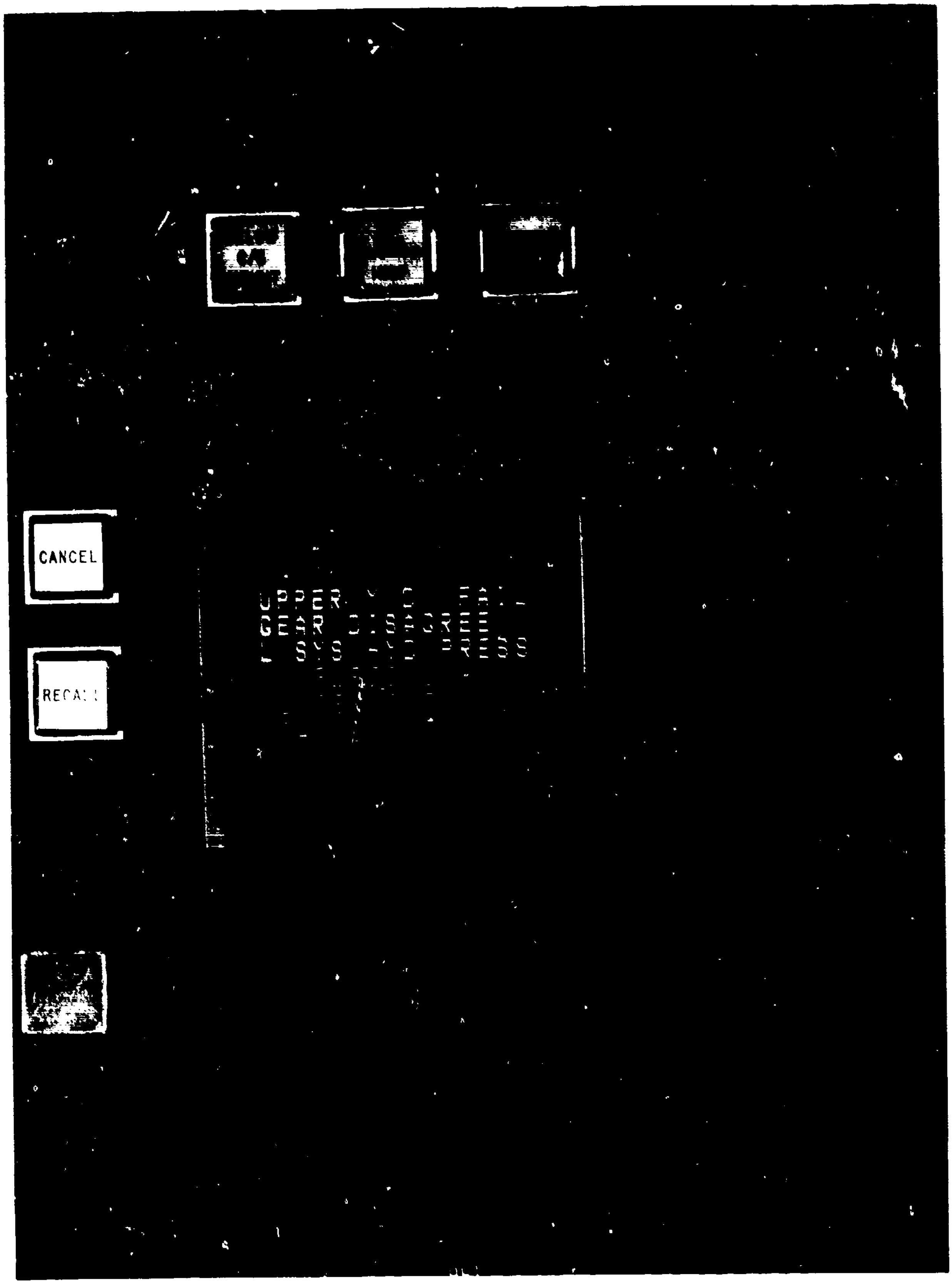


Figure 5.2.5.1.1-1. Alerts Grouped by Urgency Category—Most Recent Alert Appears at the Top of Its Group

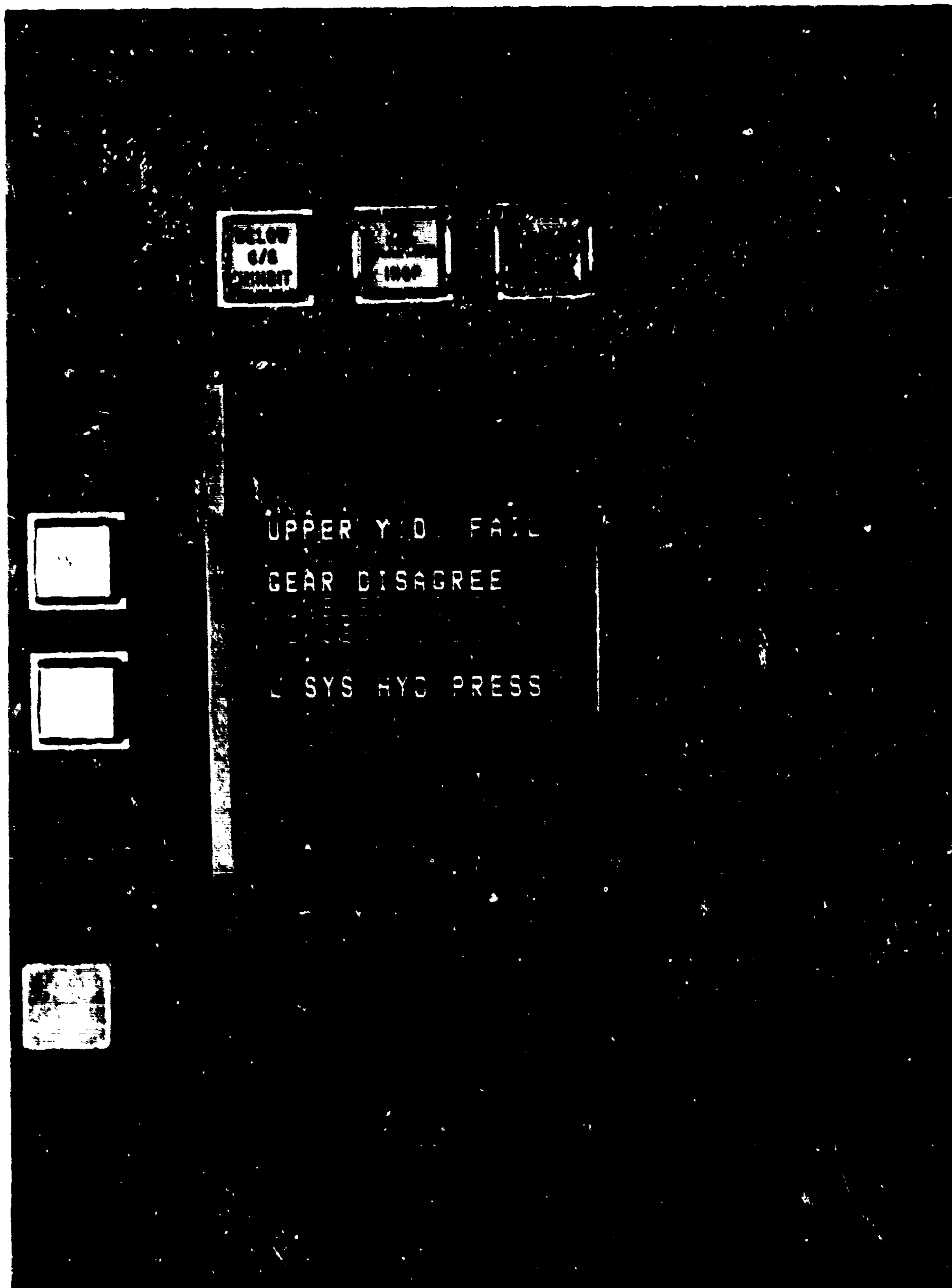


Figure 5.2.5.1.1-2. No Alert Grouping—Most Recent Alert Appears at Top of Screen

advisories) would move one line at a time until the pilot had assessed the status of his aircraft.

Finally, since system overflow is generally caused by multiple failures within a single aircraft system, the display was made to revert to a system designation (Figure 5.2.5.1.2-1); all the failing systems alerts were combined into a single message (e.g., engine #1). The bottom five lines of the display would be reserved for systems messages; a particular location on the display would be reserved for each system. When the number of displayed alerts decreased, the system would automatically revert to the normal, full message presentation.

5.2.5.1.3 ALERT CATEGORY CODING

The final area in which alternatives were investigated was the amount of visual differences that might be required to distinguish between caution and advisory messages. One concept was that if the system identifies these two levels by other means (e.g., master visual alert and tone for cautions), a large difference on the information display may not be required. A one space indentation could be used for advisories; both advisory and caution messages could be presented in the same color (amber).

An alternative argument was that each alert level should be unique and distinctive so that the pilot would waste little time identifying a new alert. This method would reduce the probability of missing advisories when a new advisory would appear between an already existing caution and advisory; it would help the pilot to more quickly determine his response to the alert, and it would also aid the pilot in assessing aircraft status especially when changing flight phases when all unresolved alerts are displayed. This necessary separation was provided in the present study by using the color blue for advisories.

A questionnaire was used to rate each concept or alternative on a number of features so that the different factors pertinent to crew alerting could be evaluated. There was also an opportunity for the pilots to present their own ideas and recommendations. The questionnaire is presented in Appendix B of this report.

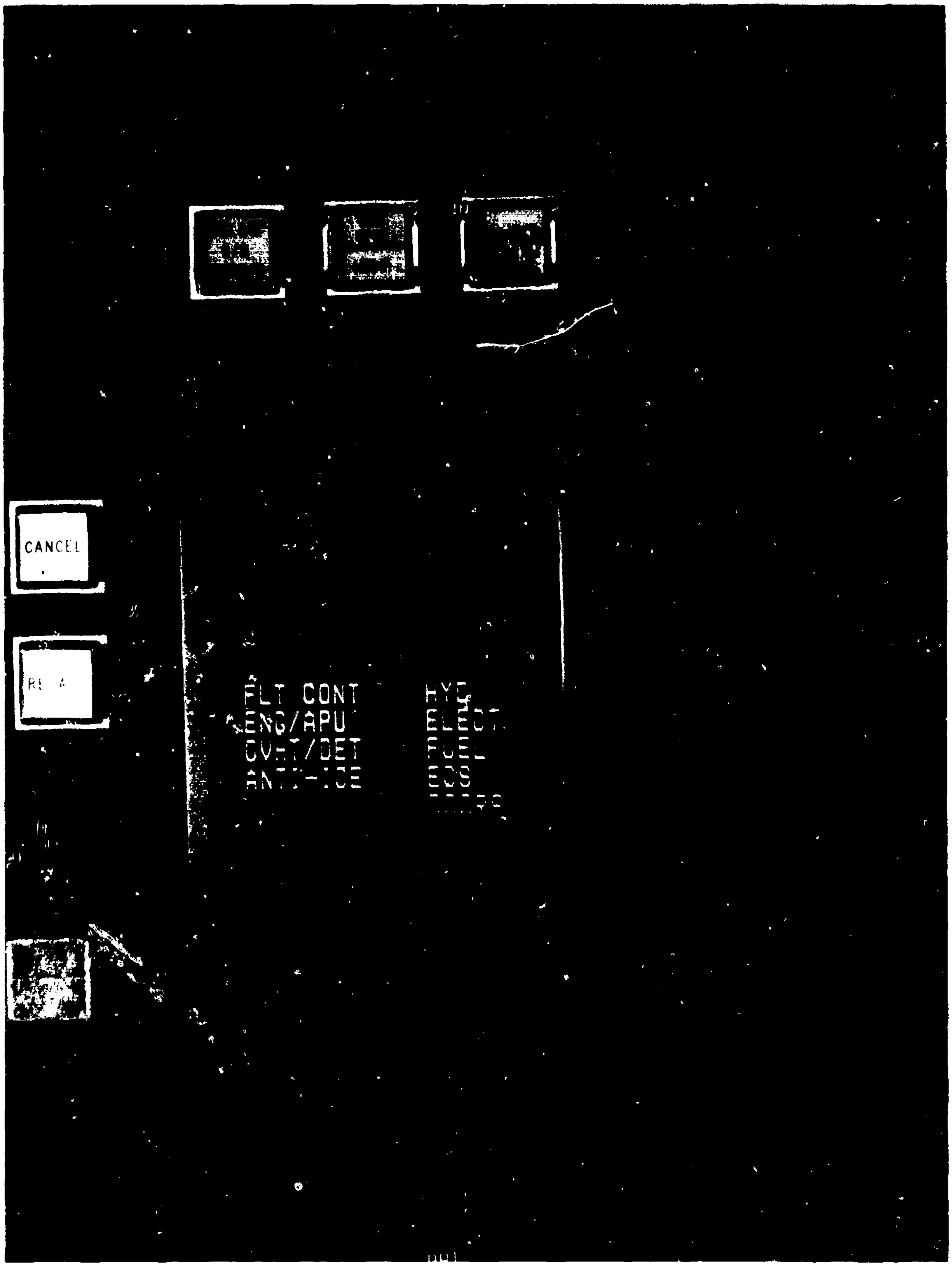


Figure 5.2.5.1.2-1. Subsystem Reversion for Overflow—All Cautions and Advisories Revert to System

Although the questions were exactly the same, the rating scale differed for the Boeing and Douglas questionnaires. The rating scale used by Boeing had anchor points only at each end and required the pilot to rate the concept on a scale from 1 to 10 (with 1 being unacceptable and 10 being excellent). Douglas used a 5 point scale with each point defined, i.e., 1 was excellent - no changes recommended, 3 was fair-minor changes recommended and 5 was unacceptable-major changes required. Both scales provided valid data about the pilot preferences, but, because of the rating scale differences the data could not be combined. However, as will be seen in the next section of this report, the correspondence between the results was quite high.

5.2.5.2 AUDITORY EVALUATION

To optimize the aural portion of the warning and caution system it was necessary to investigate the types of sounds in present use, and determine if they are appropriate for use in a new system. The questionnaire for this evaluation was administered to all 28 pilots participating in the Phase I tests, it had five sections (see Appendix B). The first section elicited pilot biographical and experience data. For sections two through five the pilots were presented with a tape recording of twelve types of of sounds used in present crew alerting systems. These sounds included: an electronic bell, mechanical bell, high wailer, low wailer, clacker, high horn, low horn, high c-chord, low c-chord, low buzzer, high chime, and low chime.

Sections three and four of the questionnaires were intended to obtain paired comparisons of the tones for attention-getting qualities and annoyance; however, because of time constraints and test scheduling problems. not all pilots received these sections. Since the amount of data obtained was not adequate to perform meaningful analysis, only sections one, two, and five are reported in the results.

Section two of the questionnaires asked the pilots to listen to each sound and identify any specific meaning they associated with it; the sounds were presented in random order to eliminate order effects.

Section three asked the pilots to listen to the sounds again (different order) and to indicate what level of urgency each suggested (warning, caution, or advisory).

Section five told the pilots to assume that a warning tone and a caution tone would precede the annunciation of all warnings and cautions. Then they were asked if there were any operational conditions which, in their opinion, would require annunciation by specific discrete tones. They were limited to a maximum of five but were not required to indicate any.

5.3 PILOT SAMPLE

The requirement for participants in the Phase I effort was that they be current transport pilots. Seven pilots, provided by the Boeing Training Group, participated in test one at the Boeing facility. These pilots were experienced instructors and also participate in flying the line with Bo g customer airlines; a summary of their experience is presented in Table 3-1; numerical entries on the right hand side of the table indicates specific experience by aircraft type and recency of the experience (1 is most recent).

Seven pilots participated in test two at the Boeing facility. Six of these were provided by the Boeing Training Group; one was from Lockheed Aircraft. All were very experienced with aircraft systems and have logged many hours flying with various airlines. Because the test data was not complete for one of the pilots, the analysis of variance was based on the data from six pilots; however, the data from the seventh pilot was evaluated and found to be comparable to that of the other six. All seven pilots completed the preference questionnaire and are included in that data. A summary of pilot experience is presented in Table 5.3-2.

A total of 16 pilots participated in tests 3 and 4 at the Douglas facility (8 pilots in each test). These pilots were engineering test pilots, and production training pilots. Engineering test pilots are typically involved with testing and evaluation of new aircraft and associated system development while production training pilots are primarly responsible for testing and delivering production aircraft as well as training pilots for customer organizations. Seven of the pilots who participated in test 4 were employed by Douglas Aircraft while one was employed by Lockheed Aircraft.

A summary of pilot experience is presented in Table 5.3-3.
A total of 28 transport pilots from the training, test and production groups of Boeing, Lockheed and Douglas participated in test 5; twelve were evaluated at the Boeing facility and 16 at the Douglas facility. A summary of the pilot experience is presented in Table 5.3-3.

The pilots surveyed at the Douglas facility were the same sample that participated in tests 3 and 4.

Table 5.3-1. Summary of Pilot Experience for Test 1

Statistical category	Pilot experience				Specific aircraft experience								
	Age	Years flying	Flight hours	Recency	707	727	737	747	DC-8	DC-9	DC-10	L-1011	A-300
Mean	47.4	28.1	11,630	1		4	3						
Standard deviation	9.0	8.3	8,490	2	1	2	3						
				3				2		1			1
Range	35-60	18-37	3,200-26,000	4	1			1			1		

Table 5.3-2. Summary of Pilot Experience for Test 2

	Pilot experience			Specific aircraft experience									
	Age	Years flying	Flight-hours	Recency	707	727	737	747	DC-8	DC-9	DC-10	L-1011	A-300
Mean	44	26	10,120	1		4	1		1			1	
				2			2				1	1	
				3		1			3				
Standard deviation	9.3	10.4	6,325	4					1	1			
				5	1		1						
				6	1								
Range	33-54	10-38	2,500-20,000										
				Total	2	5	4	0	5	1	1	2	0

Table 5.3-3. Pilot Experience for Participants in Tests 3, 4, and 5

	Pilot experience			Specific aircraft experience										
	Age	Years flying	Hours flown	Recency	707	727	737	747	DC-8	DC-9	DC-10	L-1011	A-300	Other
Boeing tested pilots														
Mean	50.8	29.9	12,960	1	1	5	2	3				1		
				2	1	3	4	3		1				
Standard deviation	5.2	5.3	5,650	3	1	4	1			1				3
				4	3		1					1		
Range	44-60	24-36	5,000-26,000	5					1					
N = 12				Total	6	12	8	6	1	2		2		3
Douglas tested pilots														
Mean	47.4		9,500	1						4	11	1		
				2						8	1			7
Range	33-58		500-21,000	3	1					2		1		12
N = 16				Total	1					14	12	2		19

5.4 PILOTS TASKS

To simulate the flight deck environment and work pattern the pilots were required to perform a tracking task using flight equipment and instrumentation. The tasks were identical for tests 1 and 2 at the Boeing facility as were the tasks for tests 3 and 4 at Douglas.

5.4.1 TESTS 1 AND 2 – FLIGHT TASK

To simulate the flight deck environment and work pattern, the pilot performed test flights of 12.5 minutes' duration in the simulator. An extremely simplistic aircraft model was used for Phase I testing; the pilots were required to track standard bar-type flight directors, respond to ATC commands and locate targets in an external visual scene.

The tracking task was presented by a Sperry AD 350A Attitude Director Indicator with standard bar-type flight directors; the flight path was generated by combining four sine waves which produced an irregular track in both pitch and roll. The equations for the pitch and roll inputs are as follows:

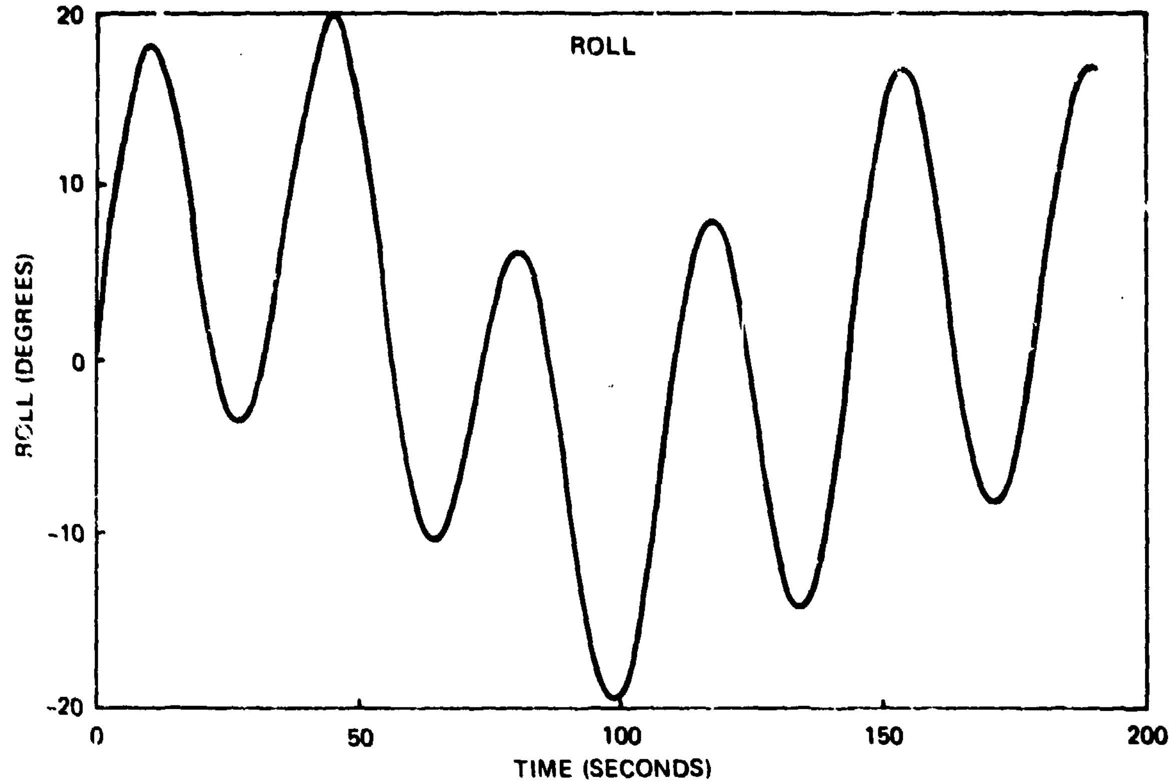
$$\text{Pitch} = K_1[W_1 + K_3\sin(D_1 \times t) + W_2\sin(D_2 \times t) + W_3\sin(D_3 \times t) + W_4\sin(D_4 \times t)]$$

Roll = Same equation with different constants

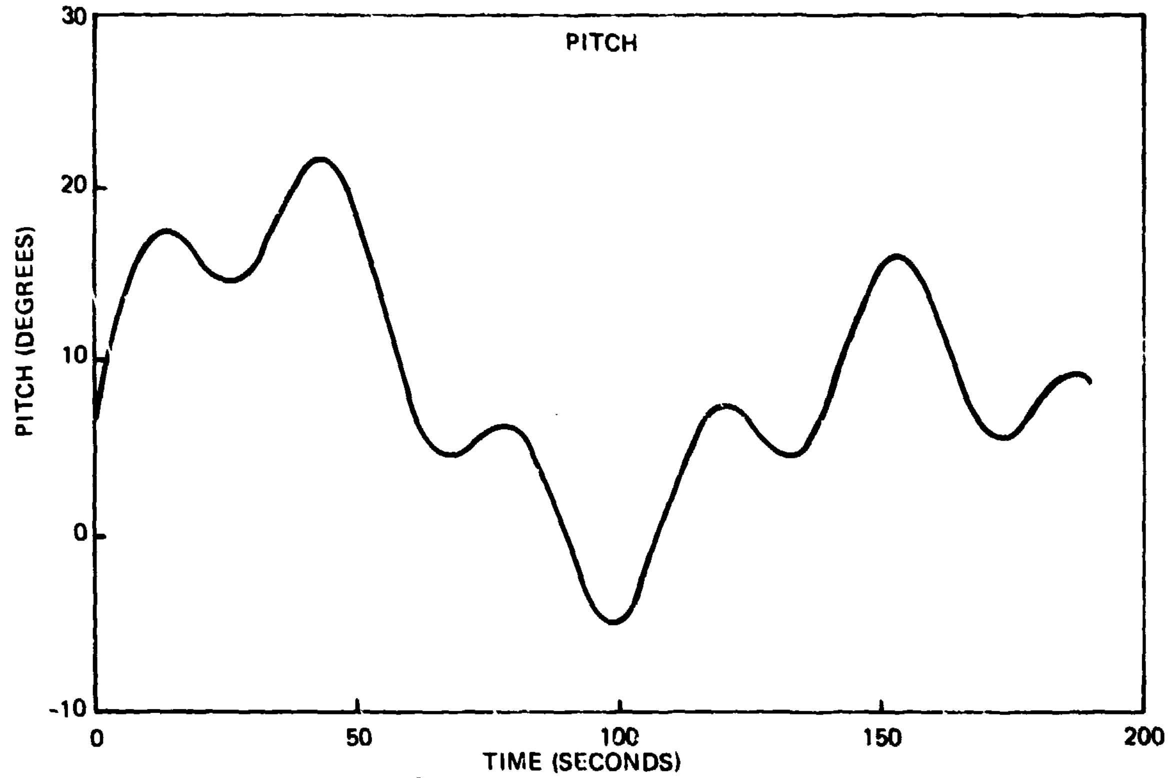
where t is the time since the start of the flight profile.

Examples of the high and low workload tracking tasks are illustrated in Figures 5.4.1-1 and 5.4.1-2. The tracking task was not designed to represent any single phase of flight but rather to provide a workload for the pilot.

To prevent the pilots from concentrating too much on the ADI and having no scan pattern at all, they were asked to respond to ATC messages which required

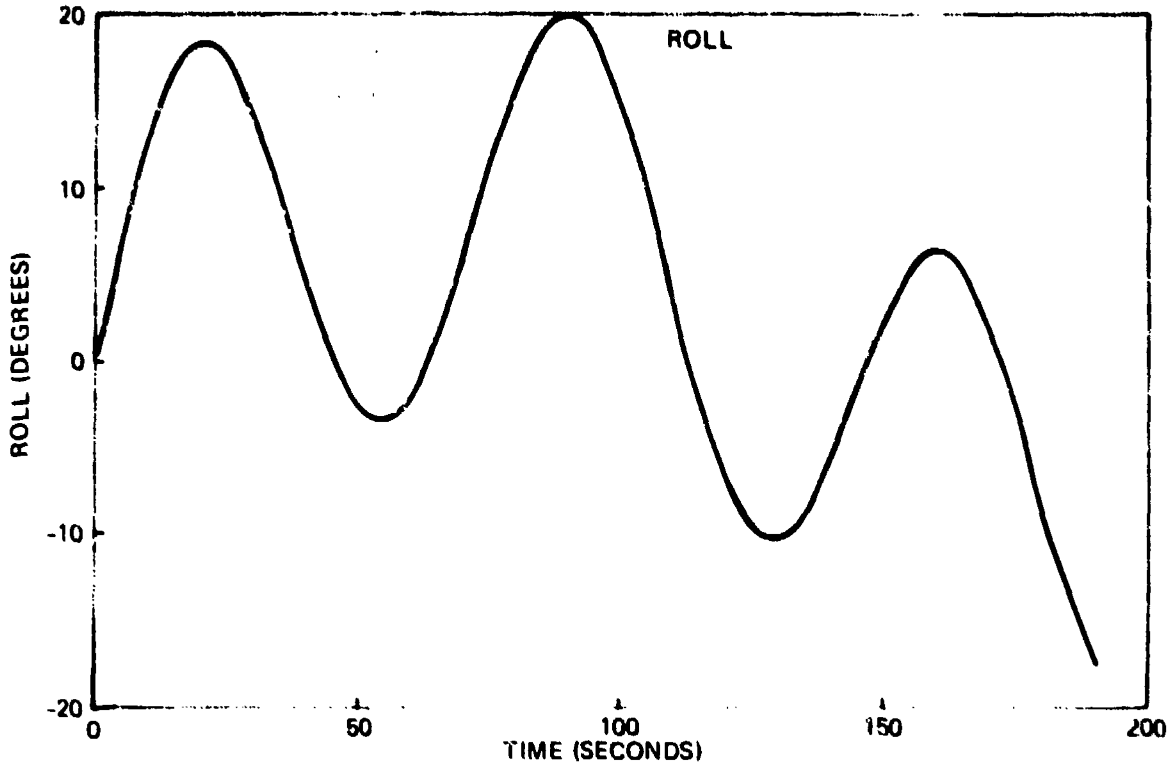


Roll = 12.5 [sin (5° x t)] + 2.5 (3.5° x t)] + 7.5 [sin (2.5° x t) + 1.25 [sin (7.5° x t)]

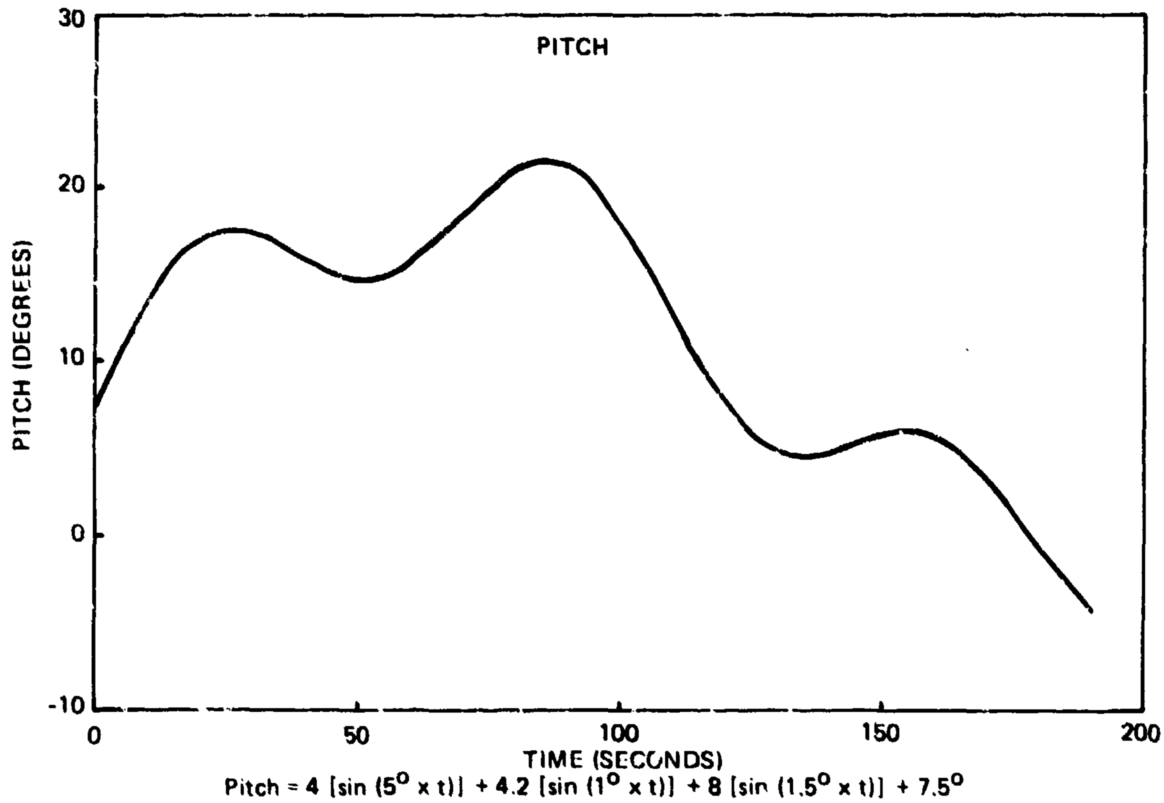


Pitch = 4 [sin (10° x t)] + 4.2 [sin (2° x t)] + 8 [sin (3° x t)] + 7.5°

Figure 5.4.1-1. Flight-Directed Path High-Workload



Roll = 12.5 [sin (2.5° x t)] + 2.5° [sin (1.75° x t)] + 7.5 [sin (1.25° x t)] + 1.25 [sin (5.25°)]



Pitch = 4 [sin (5° x t)] + 4.2 [sin (1° x t)] + 8 [sin (1.5° x t)] + 7.5°

Figure 5.4.1-2. Flight-Directed Path Low-Workload

them to hold certain airspeeds and to report heading and altitude. ATC also required the pilot to search outside the cab and locate a target on one of the three screens in front of the cab. These targets consisted of a group of split rings which were oriented in the same direction except for one ring (see Figure 5.4.1-3); each screen had some rings present for each search but there was only one target ring. The pilot's task was to locate the target ring and report its position by inputting the screen and quadrant numbers through his 10 key keyboard; Table 5.4.1-1 illustrates the numbering scheme and the order of presentation for the targets. Therefore, the flight task combined aircraft control with instrument scan and head up target detection.

One of the variables of interest in the test that had a direct effect on the flight task was the level of pilot activity (workload). Both high and low workload tasks were presented; the amount of activity was varied by changing the components of the flight task. Both high and low workload tasks followed the same flight path. Flight path excursions were programmed at a slower rate for the low workload case. The number of ATC requests for altitude and heading were doubled and the number of outside targets were increased for the high workload trials. For a further comparison of low and high workload see Table 5.4.1-2.

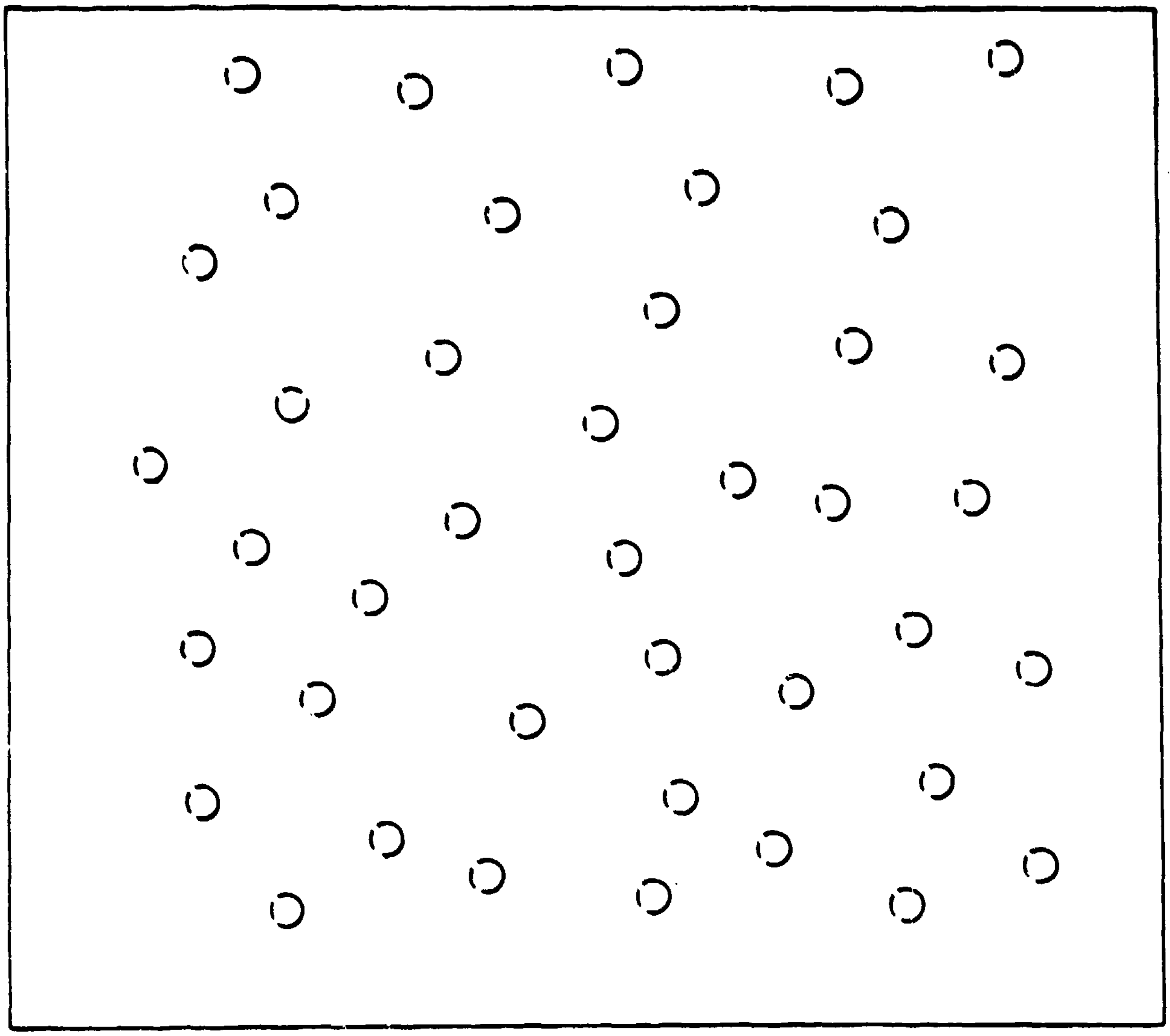
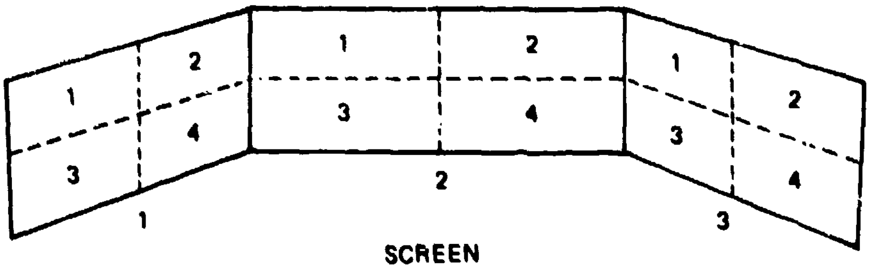


Figure 5.4.1-3. Outside Visual Search Task

Table 5.4.1-1. Answer Sheet for D-Cab Slide Sets



Slide set	Target		Slide set	Target		Slide set	Target	
	Screen	Quadrant		Screen	Quadrant		Screen	Quadrant
1	1	4	31	3	3	61	3	3
2	3	1	32	1	4	62	1	4
3	2	3	33	3	1	63	3	1
4	1	1	34	2	4	64	2	4
5	3	2	35	2	2	65	2	2
6	2	4	36	1	1	66	1	1
7	3	1	37	3	3	67	3	3
8	1	4	38	3	1	68	3	1
9	2	3	39	2	4	69	2	4
10	3	4	40	1	3	70	1	3
11	2	2	41	2	2	71	1	4
12	1	1	42	1	1	72	3	1
13	3	1	43	3	1	73	2	4
14	3	4	44	3	4	74	1	1
15	2	1	45	2	1	75	3	2
16	2	2	46	2	2	76	2	4
17	1	1	47	1	1	77	3	1
18	3	3	48	3	3	78	1	4
19	2	3	49	2	3	79	2	3
20	1	4	50	1	4	80	3	4
21	2	2	51	2	2			
22	3	4	52	3	4			
23	2	1	53	2	1			
24	1	4	54	1	4			
25	1	3	55	1	3			
26	2	1	56	2	1			
27	3	2	57	3	2			
28	1	2	58	1	2			
29	3	4	59	3	4			
30	2	3	60	2	3			

Table 5.4.1-2. Comparison of Low- and High-Workload Flight Activities

Activity	Low	High
Tracking task--number of changes in direction		
Pitch	20	42
Roll	20	42
ATC request for altitude	2	4
ATC request for heading	2	4
ATC request for traffic	3	4
Number of possible split rings	54	80
Diameter of split rings (visual angle)	0.5	0.25
Stroke width of split rings (visual angle)	0.01	0.005
Number of alerts	12	12

5.4.2 TESTS 1 AND 2 – ALERT RESPONSE TASK

When an alert was detected the pilot pressed a trigger on the control wheel and verbalized the detection. After identifying the alert he went to the response panel and pressed the switch corresponding to the system which had a problem. Table 5.4.2-1 presents all the messages programmed for the central display and Figure 5.4.2-1 identified those that were chosen. The response panel was an 18 switch panel located in the overhead and configured as seen in Figure 5.4.2-1. If the correct response was made, the message was removed from the alphanumeric display and the master alert light was extinguished. Figure 5.2.5-1 also presents the messages associated with each switch.

Table 5.4.2-1. Alert Messages

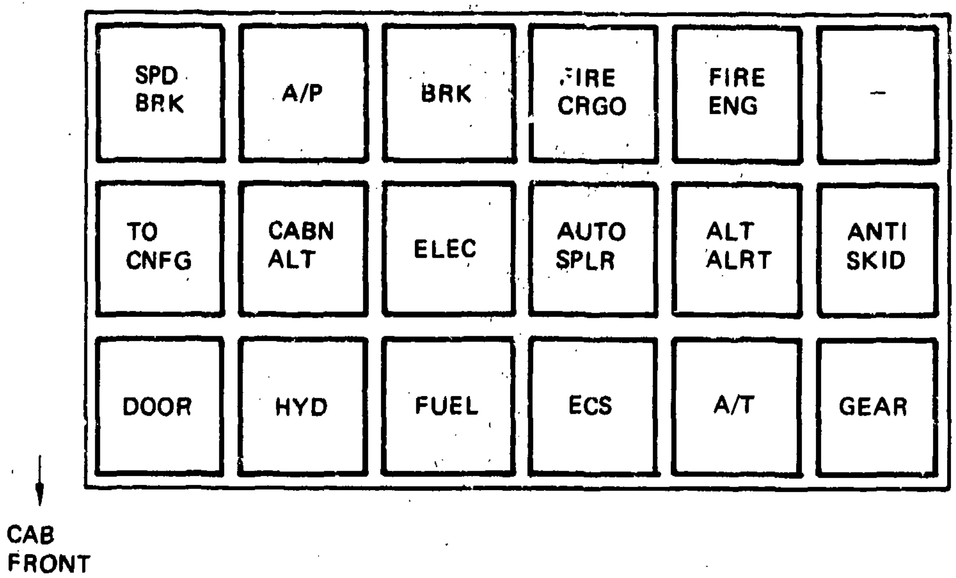
Message text	Level	Color
BLANK (for erasing a line)		Black
L ENG FIRE	W	Red
APU FIRE	W	Red
FWD CARGO FIRE	W	Red
WHEEL WELL FIRE	W	Red
TAKEOFF FLAPS	W	Red
TAKEOFF SPOILERS	W	Red
TAKEOFF STAB	W	Red
GEAR NOT DOWN	W	Red
FIRE (discrete)	W	Red
CONFIG (discrete)	W	Red
OVERSPEED (discrete)	W	Red
A/P DISC (discrete)	W	Red
STALL (discrete)	W	Red
CABIN ALT (discrete)	W	Red
PULL UP (discrete)	W	Red
VMO/MMO (discrete) (No. 11 position)	W	Red
AIRSPEED (discrete) (No. 11 position)	W	Red
PARKING BRAKE	W	Red
FLT CONTROLS	C	Amber
UPPER YD FAIL	C	Amber
L BLEED OFF	A	Blue
L GEN FIELD OFF	A	Blue
L GEN DRIVE OIL	C	Amber
L GLY BUS OFF	A	Blue
C GLY BUS OFF	A	Blue
R BLY BUS OFF	A	Blue
L UTIL BUS OFF	A	Blue
R UTIL BUS OFF	A	Blue
ELEC	C	Amber
L GEN OFF	A	Blue
L BUS TIE	A	Blue
R BUS TIE	A	Blue
GEAR DISAGREE	C	
CAT III FAULT	A	Amber
ALERT CALL	A	Blue
FWD CABIN CALL	A	Blue
L SELCAL	A	Blue
HYD	C	Amber
L SYS HYD PRESS	C	Amber
L ENG HYD PUMP	A	Blue
L ELEC HYD PUMP	A	Blue
FIRE DETECTION	C	Amber
L ENG FIRE DET	A	Blue

Table 5.4.2-1. Alert Messages (Concluded)

Message text	Level	Color
GEAR	C	Amber
ANTISKID INOP	C	Amber
L BRAKE OVHT	A	Blue
AIR COND/PRESS	C	Amber
L PACK TRIP	C	Amber
FLT DECK OVHT	A	Blue
FUEL	C	Amber
L FUEL FWD PUMP	A	Blue
DOORS	C	Amber
FWD MAIN DOOR	C	Amber
ENGINES	C	Amber
L ENG OIL PRESS	C	Amber
CAPT ATTITUDE	C	Amber
APU	C	Amber
LE FLAPS	C	Amber
ANTI-ICE	C	Amber
ADVISORY	C	Amber
FLT CONT (discrete)	W	Amber
ENG/APU (discrete)	W	Amber
OVHT/DET (discrete)	W	Amber
ANTI-ICE (discrete)	W	Amber
OVERHEAD (discrete)	W	Amber
HYD (discrete)	W	Amber
ELECT (discrete)	W	Amber
FUEL (discrete)	W	Amber
ECS (discrete)	W	Amber
DOORS (discrete)	W	Amber
ABCDEFGHIJKLMNOP		Red
QRSTUVWXYZ123456		Red
7890/-.		Red
757/767/777		Blue
TEST PATTERN		Blue
ABCDEFGHIJKLMNOP		Amber
QRSTUVWXYZ123456		Amber
7890/-.		Amber

W = Warning message
C = Caution message
A = Advisory message

81 total messages



Associated Warning and Caution Messages

Message		Switch
L ENG FIRE	(W)	FIRE ENG
TAKEOFF FLAPS	(W)	TO CNFIG
A/P DISCONNECT	(W)	A/P
GEAR NOT DOWN	(W)	GEAR
GEAR DISAGREE	(C)	GEAR
CARGO DOOR	(C)	DOOR
ANTI SKID INOP	(C)	ANTI SKID
L PACK TRIP	(C)	ECS
L BUS TIE	(A)	ELEC
L FUEL FWD PUMP	(A)	FUEL
L BRAKE OVHT	(A)	BRK
L ELEC HYD PUMP	(A)	HYD
FWD CARGO FIRE	(W)	FIRE CARGO
L GEN DRIVE OIL	(C)	ELEC
A/T DISCONNECT	(C)	A/T
AUTO SPOILER	(C)	AUTO SPLR

Note:
W = Warning
C = Caution
A = Advisory

Figure 5.4.2-1. Overhead Response Panel

5.4.3 TESTS 3 AND 4 —FLIGHT TASK

The test paradigm for test 3 and 4 was based on a simulated final approach scenario. The VFR approach and landing segment was selected for experimental trials because it represents a highly time-critical phase of flight in which rapid and accurate decisions are essential. This flight regime also provides an appropriate environment for evaluation of distracting or masking effects since final approach maneuvers are generally associated with high levels of visual and auditory task loading.

The basic elements of the part-task simulation and a typical event sequence are illustrated in Figure 5.4.3-1. Each experimental trial was initiated with a landing clearance from the tower controller. During the approach, the pilot heard a series of ATC communications presented against a typical cockpit noise background. At some point prior to reaching the middle marker, a specific ATC advisory message was directed to "DACO 891". This call sign represents the aircraft designation assigned to the simulator cockpit. The particular profile illustrated in Figure 5.4.3-1 shows an ATC message that is concurrent with an alert message (simultaneous onset at the time reference T_A).

During the simulated approach sequences, test subjects were required to perform the following basic tasks:

1. Maintain safe and accurate control of the aircraft from the outer marker to termination of the test trial,

2. Monitor ATC transmissions and respond verbally to selected messages from the controller,

3. Identify and acknowledge auditory alert messages generated by the central aural warning system.

The pilot employed Head-up Display (HUD) flight director cues as the primary visual reference for control of the aircraft. No active head-down flight instruments were provided. The basic HUD symbology available for guidance during approach and landing is illustrated in Figure 5.4.3-2. The essential

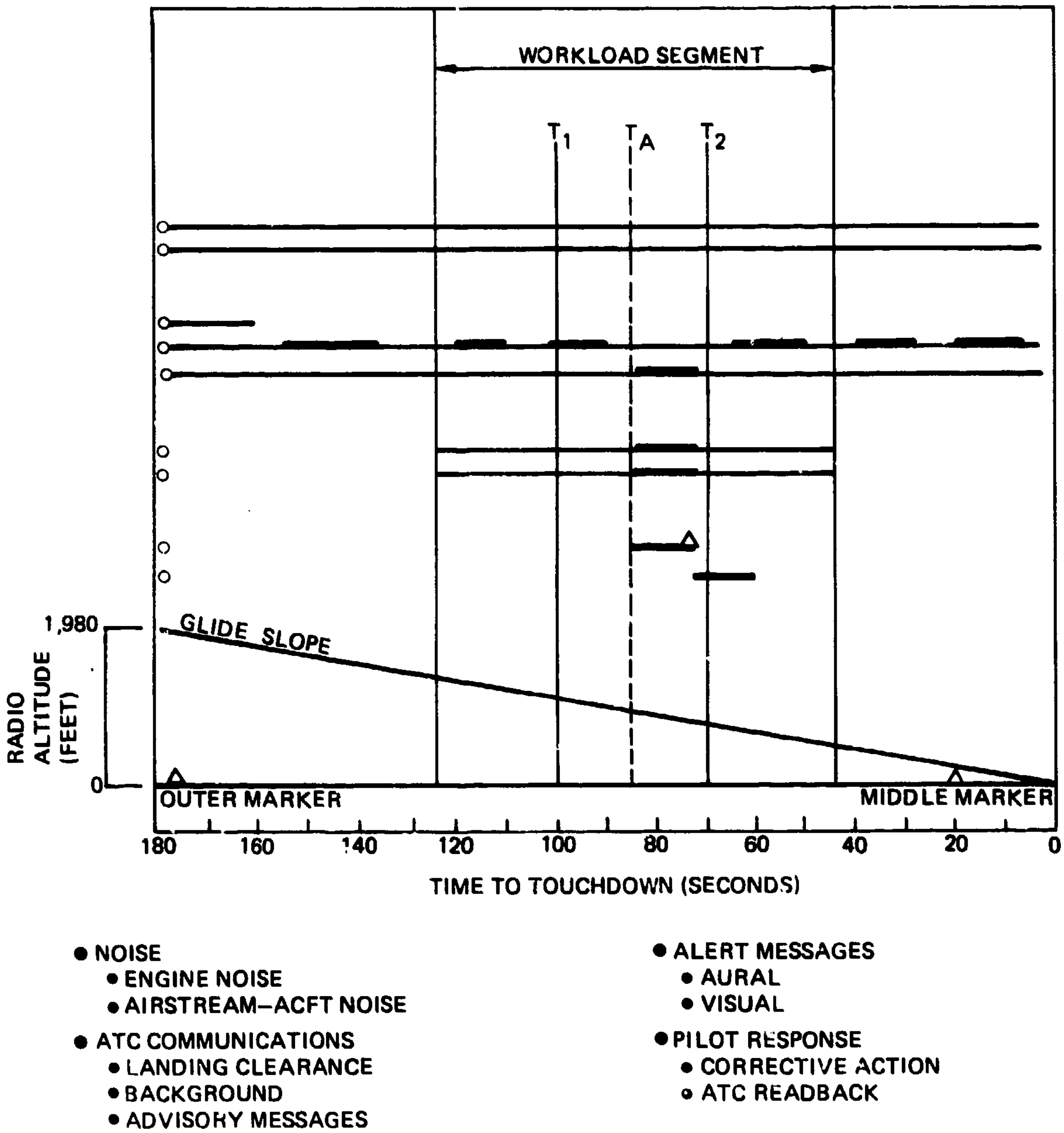


Figure 5.4.3-1. Basic Test Scenario

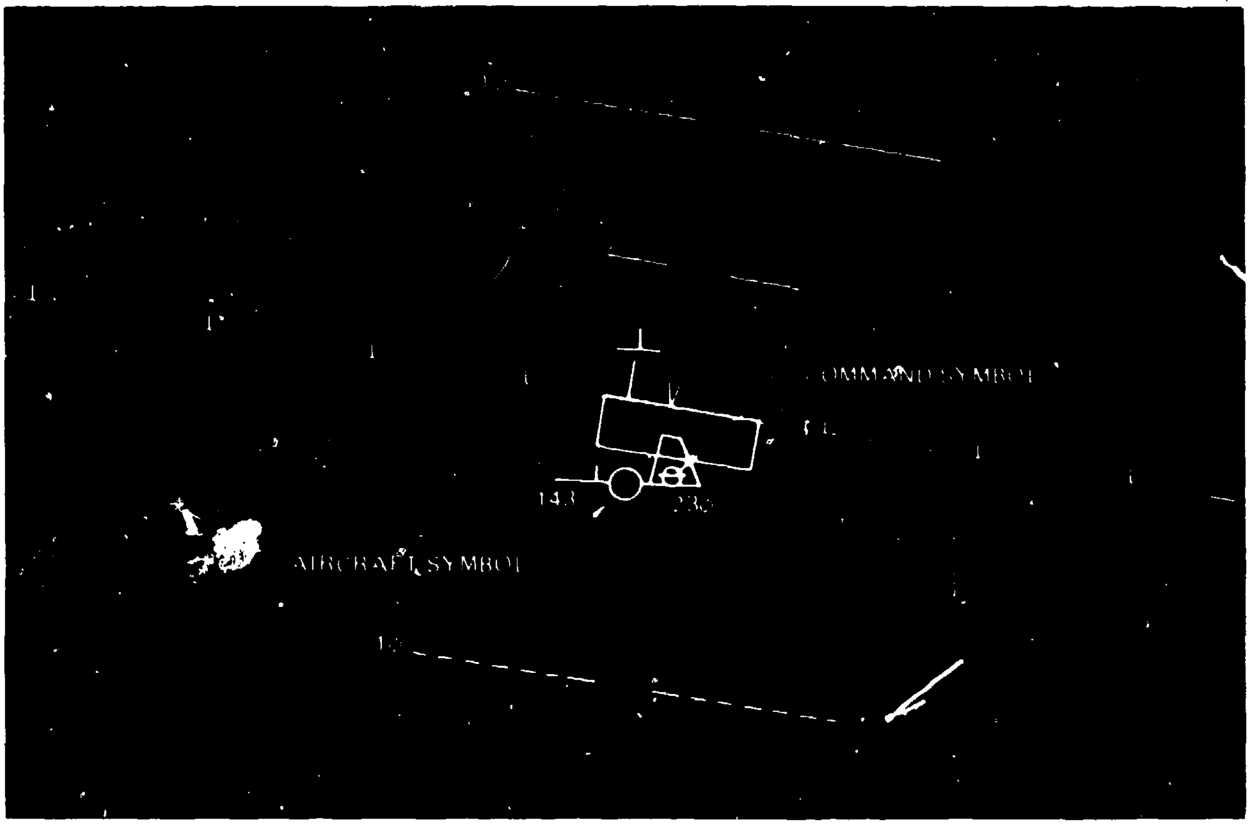


Figure . 3-2. HUD Symbology Used for Tests 3 and 4

3

components of the HUD format used in the test trials are the flight director command symbol and the aircraft symbol. The pilot's task was to maintain the aircraft symbol centered over the flight director target as accurately as possible through final descent, flare and touchdown. The external visual scene was represented on an Advent projection screen at a distance of approximately fifteen feet.

Control of the aircraft was exercised by means of a side-stick controller. Although no rudder pedals were installed, a yaw damper was simulated in the flight task program to augment lateral stability. Autothrottles were engaged at all times and generally maintained airspeed within two knots of the command value. The handling qualities of the simulator were patterned after those of a DC-10-30.

Each approach was initiated at an altitude of 1980 feet on a 3^{o} glideslope with a target airspeed of 142 knots. The simulated wind profile consisted of random turbulence introduced gradually during the first 20 seconds of the approach and maintained at a moderately high level through the remainder of the test run. The simulation was terminated at the point of touchdown and the simulator was reset for the next trial.

5.4.4 TESTS 3 AND 4 – ATC RECOGNITION TASK

The air traffic control simulation consisted of two channels of tape-recorded ATC communications which were presented against a representative ambient noise background. The ATC recordings were developed in cooperation with FAA personnel at the Los Angeles International Airport (LAX) control tower. ATC traffic recorded on channel 1 of the tape, was designed to simulate messages to other aircraft under local control and general "chatter" that is commonly heard during terminal area flight maneuvers. Recordings of experienced male and female controllers were obtained by monitoring transmission during normal VFR operations at the LAX facility. The frequency and content of communications were typical of high density traffic periods at major airports. Communications recorded on channel 2 were specific advisory messages and instructions addressed to "DACO 891". The controllers were asked to make recordings of eight different messages using their normal speech rate and

voice inflection. Each advisory message was later dubbed onto separate tape cassettes to facilitate playback at specific time reference points for the test trials. During the tape transfer process, the peak loudness levels for male and female controller voices were adjusted to yield equivalent average weighted sound level measurements. The specific ATC transmissions, message durations, and sound levels recorded in the cockpit are listed in Appendix C.

The background noise simulation was designed to represent the ambient noise characteristics of a commercial aircraft cockpit. The original sound track was recorded during the final approach segment of a DC-10-30 production test flight using a calibrated microphone mounted 40 inches above the cockpit floor, directly between the Captain and First Officer. Since the acoustic characteristics of the DETAC cockpit shell differ significantly from those of an actual flight deck, some modification of the recordings was necessary to duplicate the appropriate spectral contour at the pilot's ear position. This was accomplished by re-recording the noise track through the use of one-third octave band-shaping equipment and a real-time analyzer. The loudness level for ambient noise in the cockpit was held constant at 73 dBA.

The onset of the ambient noise recording was coincident with initialization of the test trial. Shortly after crossing the outer marker, the pilot was given a clearance to land on Runway 25 left. This was followed by a series of radio transmissions, spaced at random intervals and continuing through the remainder of the approach profile. An ATC advisory message directed to "DACO 891" was presented at a pre-determined point during each test run. On 50% of the trials (high auditory workload condition) the ATC message was annunciated concurrently with an auditory alert. On these trials, the onset of alert and ATC messages was essentially simultaneous, although minor variations of up to 250 msec were observed in the ATC onset times due to the variability of the cassette rewind mechanism. Upon hearing an ATC advisory message directed to "DACO 891", the pilot's task was to attempt to read back the essential parts of the message. If the message was unintelligible or the pilot failed to hear the entire communications, he was instructed to request a message repeat. The accuracy of the pilot's verbal response was noted by the experimenter and recorded on tape for subsequent verification.

5.4.5 TESTS 3 AND 4 —ALERT RECOGNITION TASK

During the course of the experimental trials, test subjects were required to acknowledge aural alerts by pressing one of eight labelled push-buttons on the overhead panel. For purposes of the alerting system tests, pressing the appropriate button represented initiation of a corrective action and indicated that the pilot had identified and correctly interpreted the fault message. The arrangement of the response panel is shown in Figure 5.4.5-1. The switch matrix consisted of eight back-lighted push buttons with white legends. The switch legends were illuminated at all times for reading purposes. Each of these switches corresponded to one of eight possible alert indications representing system failures or deviations from normal aircraft configuration. Although the test subjects were not required to differentiate between alert priority levels in making their responses, the alert message set included four warning and four caution level failures.

On each test trial, a failure message was introduced at a pre-programmed time reference point designated T_A. To facilitate isolation of potential distracting effects resulting from the alert onset, specific values of T_A were allowed to vary within a 120 second time interval identified in Figure 5.4.3-1 as the "Workload Segment". Based on pilot performance during preliminary trials, it was determined that these limits defined a range within which the level of difficulty of the flight director task remains relatively stable. It was necessary to restrict alert onset time to this range to assure that any decrements in tracking accuracy could be attributed to the onset of alert messages and to avoid contamination of the post-alert performance measures by systematic changes in difficulty of the loading task over time.

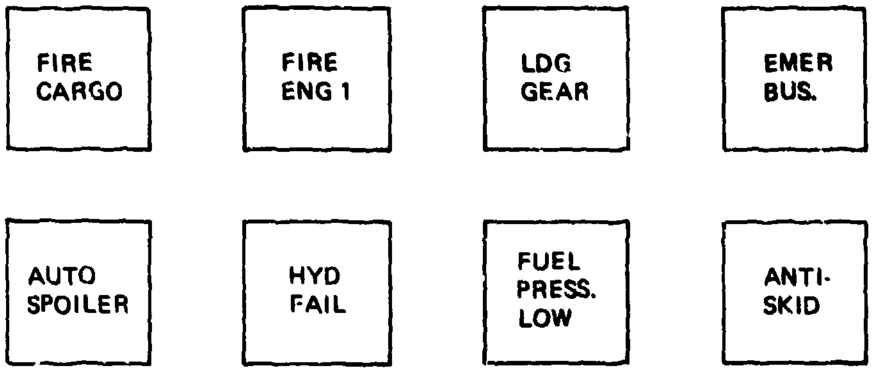


Figure 5.4.5-1. Overhead Response Panel

5.5 TEST PROCEDURES

5.5.1 TESTS 1 AND 2

All variables not tested were held constant or controlled to avoid biasing or confounding the results: Aircraft ambient noise of approximately 75 dB was presented during the flight task to mask the uncontrolled noise that may be occurring around the cab; ambient lighting was relatively low (i.e., 100 fl) to permit the use of the outside scene; ATC communications were presented 5-10 dB above the ambient noise and held constant for all trials; visual message contrast was set at the same predetermined nominal figure for each flight; interior lighting was controlled and did not vary during trials; all the pilots received the same instructions to minimize experimenter bias (see appendix A).

Each test flight was twelve and a half minutes in length and contained twelve alert messages, four from each priority level. The alerts were presented on a schedule around the even minutes; however, to prevent the pilots from learning the timing of the events a 40 second time period around each even minute was allocated for the alert (minute $\pm$ 20 seconds). The time was chosen at random and 12 different time scenarios were developed (see Table 5.5.1-1). The only restriction placed on the selection of the deviation was that the last alert could not occur more than 5 seconds past the 11 minute mark so that the pilot had at least 25 seconds to respond before the end of the flight.

To reduce the possibility of influencing the data because of the way the alert messages were presented, 12 random orderings were developed for presenting the alerts. Not only were different orders presented, 4 substitute alerts were also included in 5 of the orders to further reduce any bias resulting from the pilot learning the message set. The 12 alert scenarios and the message set are presented in Table 5.5.1-2. The time and alert scenarios were randomly combined for each test.

Whenever task performance is measured under several different treatment conditions, learning or fatigue may affect performance on later trials. Care must be taken in such cases to prevent these carry-over effects from

Table 5.5.1-1. Time Scenarios for Alerts

• 12.5-min flight with alerts at 1 min ±20 sec intervals

Presentation time*	Scenario number										
	1	2	3	4	5	6	7	8	9	10	11
1	+5**	-10	-6	-8	-6	+7	+5	+15	-2	+19	+17
2	-14	-16	+13	+7	+1	+17	-14	-16	+11	-20	+11
3	+11	-20	-8	+5	-10	-14	-16	-2	-10	-16	-6
4	-16	+15	-12	+15	+11	+13	+11	+5	+9	-14	-16
5	+1	+3	+1	-10	+15	-4	-6	+19	-8	-8	+15
6	-18	+7	+9	+1	-8	-12	+15	+7	+1	-2	+13
7	+17	-8	+17	+9	+7	-6	+13	-8	+13	+9	+1
8	-20	+19	+5	-4	-2	+19	+3	+3	-18	-18	-18
9	-10	+13	-10	-6	-14	+1	+1	-4	+3	-10	+7
10	+15	+17	+11	-20	-4	+3	+17	-12	-16	-4	+19
11	-2	+5	-18	+11	+13	-20	+19	-6	-12	+5	+5
12	-6	+1	-14	+1	-20	-10	-2	+1	-14	-6	+3
12.5	End										

*Time in minutes after start

**Column figures represent adjustment (in seconds) to presentation time

Table 5.5.1-2. Alert Scenario

Presentation order	Scenario number* 1	2	3	4	5	6	7	8	9	10	11	12
1	11	5	4	7	1	12	3	8	10	2	9	6
2	2	8	7**	10	4	3	6*	11	1	5	12	9
3	10	4	3	6	12	11	2	7	9	1*	8	5
4	9	3	2	5	11	10	1*	6	8	12	7	4
5	8	2	1*	4	10	9	12	5	7	11	6	3
6	5	11	10	1	7*	6	9	2	4	8*	3	12
7	12	6	5	8	2	1	4	9	11	3	10	7
8	7	1	12	3	9	8	11	4	6	10	5	2
9	3	9	8*	11	5	4	7*	12	2	6*	1	10
10	1	7	6*	9	3	2	5	10	12	4	11	8
11	6	12	4	2	8*	7	10	3	5	9	4	1
12	4	10	9	12	6*	5	8*	1	3	7*	2	11

Legend:

Alert number*		
1	L ENG FIRE	(w)
2	TAKEOFF FLAPS	(w)
3	A/P DISCONNECT	(w)
4	GEAR NOT DOWN	(w)
5	GEAR DISAGREE	(c)
6	CARGO DOOR	(c)
7	ANTISKID INOP	(c)
8	L PACK TRIP	(a)
9	L BUS TIE	(a)
10	L FUEL FWD PUMP	(a)
11	L BRAKE OVHT	(a)
12	L ELEC HYD PUMP	(a)

HOLDOVER ALERTS	
UPPER YD FAIL	(c)
L ENG OIL PRESSURE	(c)
CAT II FAULT	(a)
CABIN ALTITUDE	(w)

SUBSTITUTE ALERTS	(**)
1** FWD CARGO FIRE	(w)
6** AUTO SPOILER	(c)
7** L GEN DRIVE OIL	(c)
8** A/T DISCONNECT	(c)

Notes:

*Numbers in the table correspond to alerts.
**Use substitute messages.

(a) Advisory
(c) Caution
(w) Warning

differentially affecting the performance measures for the different treatment conditions. It should be noted, therefore, that the order in which the pilots received the experimental treatment was also randomly assigned to prevent order biases from confounding the results (see Table 5.5.1-3 and 5.5.1-4). Immediatly prior to each test flight the pilot was informed what specific configuration of the alerting system he would be using during that flight. (e.g., dual master).

A daily test schedule for the tests 1 and 2 are presented in Table 5.5.1-5 and 5.5.1-6. A total of seven pilots were tested with one day required for each pilot to complete test 1 and one and one-half days for test 2.

The test session began with an introduction to the Developmental Cab. This introduction included a brief description and walk-through of the simulator facility. Following the introduction, the pilot was briefed on the "aircraft's" operational characteristics and the flight task, including a description of the warning/caution system that was to be used (see appendix D); he was also instructed in the procedures to be used during the tests. The briefing took place in the cab so that the actual systems could be used for illustration.

After the briefing period the pilot was permitted to fly the simulator to become familiar with the flight task. After 3 practice flights (12.5 minutes each) the test flights began. Each pilot performed 12 flights, each twelve and one-half minutes long.

Each test flight began with the aircraft flying straight and level at 110 knots and with the flight director bars centered. The pilot was required to follow the flight director, to hold speed, and to respond to alerts to ATC communications by identifying traffic and reporting flight parameters.

At the end of the test trials, each pilot completed a questionnaire concerning his subjective evaluation of the alerting concepts and a number of questions were asked about the different formats and treatment (see appendix B). Pictorial representations of the alerting concepts were available to help the pilots recall the test conditions; the test conductor was present during the

Table 5.5.1-3. Presentation Order for Test Flight in Test 1

Attention format		Single master						Dual master—no flash						Dual master with flash					
Display format		All separate		Warnings separate		No separation		All separate		Warnings separate		No separation		All separate		Warnings separate		No separation	
Workload		High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
Subjects	1	17	18	14	13	1	2	16	15	5	6	8	7	11	12	10	9	3	4
	2	12	11	7	8	14	13	9	10	18	17	1	2	6	5	3	4	16	15
	3	9	10	6	5	11	12	8	7	15	16	18	17	3	4	2	1	13	14
	4	4	3	17	18	6	5	1	2	10	9	11	12	16	15	13	14	8	7
	5	13	14	10	9	15	16	12	11	1	2	4	3	7	8	6	5	17	18
	6	8	7	3	4	10	9	5	6	14	13	15	16	2	1	17	18	12	11
	7	15	16	12	11	17	18	14	13	3	4	6	5	9	10	8	7	1	2

Table 5.5.1-4. Presentation Order for Test Flight in Test 2

Display location	Single center								Warning pilot caution center								Both pilots							
Attenson	Yes				No				Yes				No				Yes				No			
Message flash	Yes		No		Yes		No		Yes		No		Yes		No		Yes		No		Yes		No	
Workload	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
Subjects 1	9	10	16	15	5	6	12	11	14	13	19	20	24	23	3	4	22	21	7	8	18	17	1	2
2	16	15	21	22	12	11	17	18	19	20	2	1	5	6	10	9	3	4	14	13	23	24	3	4
3	7	8	3	4	10	9	12	11	17	18	22	21	2	1	20	19	13	14	5	6	16	15	24	23
4	6	5	11	12	2	1	7	8	9	10	16	15	19	20	24	23	17	18	4	3	13	14	21	22
5	3	4	10	9	23	24	6	5	8	7	13	14	18	17	21	22	16	15	1	2	12	11	20	19
6	22	21	3	4	18	17	23	24	1	2	8	7	11	12	16	15	9	10	20	19	5	6	13	14
7	11	12	18	17	7	8	14	13	16	15	21	22	2	1	5	6	24	23	9	10	20	19	4	3
8	2	1	7	8	22	21	3	4	5	6	12	11	15	16	20	19	13	14	24	23	9	10	17	18

Table 5.5.1-5. Daily Test Schedule for Test 1

Test hours	Action
0000 to 0100	Bring simulator on line and perform checkout
0100 to 0110	Introduction to simulator facility
0110 to 0125	Briefing on flight task and alerting systems
0125 to 0200	Practice flight
0200 to 0300	Test flights 1 through 4
0300 to 0315	Break
0315 to 0445	Test flights 5 through 10
0445 to 0530	Break
0530 to 0730	Test flights 11 through 18
0730 to 0800	Debriefing

Table 5.5.1-6. Daily Test Schedule for Test 2 (2-Day Test)

Test hours	Action
Day one	
0000 to 0100	Bring simulator on line and perform checkout
0100 to 0115	Introduction to simulator facility
0115 to 0145	Briefing on flight task and alerting systems
0145 to 0200	Practice flight
0200 to 0300	Test flights 1 to 4
0300 to 0315	Break
0315 to 0445	Test flights 5 to 10
0445 to 0530	Break
0530 to 0730	Test flights 11 to 18
0730 to 0800	End-of-day debriefing
Day two	
0000 to 0100	Bring simulator on line and perform checkout
0100 to 0130	Retraining session
0130 to 0300	Test flights 19 to 24
0300 to 0345	Final debriefing
0400	Begin next subject

debriefing to answer any questions and to assure that all items were answered completely. Each pilot also completed a standard biographical data sheet concerning flight experience.

5.5.2 TESTS 3 AND 4

Primary experimental investigation took place in the DETAC facility at Douglas Aircraft. Three hours of simulation time were dedicated to the study on a daily basis to insure the successful completion of all procedural activities.

In designing an appropriate counterbalancing scheme for a number of objectives had to be met. The first was that each experimental condition (or cell) be used as a different trial for each subject. A second objective was for the ATC and voice alert messages to be paired randomly, for each cell, meaning that a particular ATC message would not be paired with the same voice alert message for every pilot. A third objective was to insure that the aural alerts would be distributed evenly over the workload segment. Using eight subjects, it was possible to meet all of the above objectives by using balanced square designs. For the subject runs (Objective 1), two balanced squares were used. Subjects 1 to 4 received the word messages first and then the sentence messages. The messages were administered to subjects 5-8 in the opposite fashion. The second objective was met in a similar manner. Two balanced squares were used in arranging the aural alert messages (one each for the word and sentence formats) and two inverted balanced squares were used for the ATC messages (one each for male and female message sets). It was also possible to meet Objective 3 using a balanced square design. Since the aural alerts could be introduced at any one of eight different times, a number from one to eight was assigned to each time. The odd numbers represented times in the first half of the workload segment while the even numbers corresponded to times in the second half of the workload segment. Using a balanced square with this configuration, it was possible to assure that for every subject, one half of the aural alerts occurred in the first half of the workload segment and the other half in the second, each subject hearing a different set of alerts in each segment. It should be mentioned that aural alerts were introduced as a function of ATC messages. The time that each aural alert was introduced depended on the time that the corresponding ATC message was introduced. Again, two balanced squares were used to assure an equal and balanced distribution of ATC onset times.

For test 4, a number of additional objectives had to be met. There were four alerting modes as well as a control condition where no alerts were introduced. Each mode contained four experimental conditions (or cells). Each of the four experimental conditions within the four alerting modes represented a unique combination of ATC and alert messages presented at specific times within a pre-defined workload segment. By utilizing a balanced square design, it was possible to maintain this unique configuration. A second objective was for the Air Traffic Control (ATC) messages and the alert messages to be paired with the same alert message for every subject. This was achieved by using two balanced squares for the alert messages and two inverted balanced squares for the ATC messages. A third objective was to insure that the alert messages would be distributed evenly over the workload segment; a balanced square was employed to meet this requirement.

A daily test schedule for tests 3 and 4 is presented in Table 5.5.2-1. A total of 16 pilots were tested with four hours required for each pilot to complete the test program.

Before entering the DETAC facility, pilots were asked to complete an alerting tone evaluation questionnaire (See Appendix B). This questionnaire was administered in conjunction with the playing of a cassette tape containing a number of alerting tones currently being used in commercial aircraft. Prior to administration of the questionnaire, pilots were given a general description of the rationale behind this procedure. Pilots were encouraged to voice questions or comments at any time during this session, which lasted approximately 30-45 minutes. The alerting tone evaluation took place in a presentation room adjacent to the acoustics laboratory at the Douglas - Long Beach facility. This room, which was used for all test subjects, is considered to be acoustically "dead" and employs two overhead speakers positioned approximately 4 feet above and 12 feet to either side of the pilot. All tones were presented at an average peak sound level of 85 $\pm$ 3 dBA. Ambient sound level in the room was approximately 75 dBA.

Upon entering the DETAC facility, several operational procedures were required before actual testing could begin. In the interest of time, the pilot was briefed by one experimenter while another experimenter initialized and verified the various system components.

Table 5.5.2-1 Daily Test Schedule for Tests 3 and 4

Activity	Cumulative hours
Pretest alerting tone evaluation	0000 to 0030
System startup and checkout procedure	0030 to 0100
Test briefing and instructions	0030 to 0100
• Basis tasks	
• Active displays and controls	
• Flight task description	
• Voice message format	
Practice trials	0100 to 0130
• Familiarization with HUD symbology dynamics	
• Familiarization with low-workload flight task	
• Familiarization with nigh-workload flight task	
• Familiarization with high-workload flight task	
• Various voice alerts and ATC communications	
Test trials	
• Trials 1 through 8	0130 to 0300
• Break	
• Trials 9 through 16	
Debriefing session	0300 to 0345
• Debriefing checklist	
• Overflow logic evaluation	
• Informal discussion	

The briefing/instruction period was carried on interactively with the experimenter explaining each component of the test scenario, while allowing the pilot to interject comments or questions when desired. Following an introduction to the DETAC facility, the pilot was briefed on the "aircraft's" operational characteristics and the flight task, including a description of the warning/caution system (see Appendix D).

The pilot was informed as to the three basic tasks to be carried out during each trial. The visual tracking task involved a series of approaches through moderate to high turbulence using the head-up display with flight director cues (projected on the Advent projection screen). The second task consisted of monitoring ATC communications. Pilots were instructed to read back transmissions directed to their aircraft. Finally, responses to voice warning messages were to be made by pressing the appropriate button on the overhead panel.

The cockpit contained no active head-down displays. Three separate speakers used for the aural presentations were located as follows:

ATC - Overhead

Voice alerts - Side panel

Ambient noise - Cockpit floor

Selected controls were also activated for the test. The side stick controller was operational and pilots were advised as to its relatively high sensitivity. Other active controls included the CAWS cancel switch as well as the overhead response switches, the depression of which represented corrective action in response to a fault message.

Next the flight task requirements were explained in detail. Specific information for the visual tracking task included a HUD symbology briefing, initial approach speed and altitude, and specific turbulence characteristics. This was followed by an elaboration on the ATC recognition task which included directions for responding to an ATC runway clearance message, the aircraft

call letters and specific responses that would be appropriate. The voice alert task was then described more fully, including an explanation of the voice characteristics, location of the response panel, alert duration and the appropriate message response sequence.

Following the flight task description, the specific voice message formats were reviewed with particular attention being given to format changes the pilots would experience during the test session. At this point, the alerting tones and voice messages were annunciated and the pilots were asked to acknowledge them by depressing the appropriate switch on the overhead response panel. This period of time served to familiarize pilots with the response panel nomenclature and layout, as well as giving them an opportunity to hear each alerting tone and voice alert message.

Before participating in any test trials, each pilot made a series of practice runs. Since the purpose of the practice trials was to bring the pilots to a pre-determined performance standard, the length of this period varied from pilot to pilot.

The first practice trial consisted of a high turbulence flight with the autopilot engaged. This trial was used primarily to demonstrate the dynamics of the HUD symbology. No alert messages or ATC communications were introduced at this time. During the second practice trial, the autopilot was disengaged and no turbulence was introduced. This trial served to familiarize the pilot with the side stick controller and the nature of the tracking task. As with the first practice trial, no voice alerts or ATC communications were introduced. At the point of touchdown, various performance parameters (e.g., sink rate, lateral and longitudinal dispersion) were projected onto the Advent projection screen. It was explained to the pilot that these would be used as an aid in determining when a stable level of performance was reached. Following this, several trials were run where high turbulence was introduced and the autopilot was disengaged. These trials helped to familiarize the pilot with the tracking task under high workload conditions and were repeated until the pilot reached a stable performance level. As with the first two practice trials, no voice alerts or ATC communications were introduced. When it was clear that the pilot was performing consistently, a series of trials

were run where both voice alerts and ATC communications were presented. These practice trials served to familiarize the pilots with all tasks that would be part of the test scenario.

Prior to beginning the test trials, a number of final system checks were made to insure a relatively trouble-free test session. The data recording system was initialized via the cockpit control panel. The video/audio recording equipment, the ATC communication and ambient noise track units were rechecked and initialized.

The test session consisted of 16 trials, each lasting approximately three and one-half to four minutes. System resets between each trial consumed an additional 30 to 60 seconds so that actual testing time ranged from 60 to 90 minutes. A brief rest period was taken after the eighth trial in an effort to reduce the effects of fatigue. Additional time was allowed for pilots to interact with the experimenters between trials, as this happened periodically. With all aspects of the test session included, the total elapsed time spent with one pilot in the DETAC facility was approximately 3 hours.

Upon completion of the test trials, pilots were invited to a debriefing session in an area outside the DETAC facility. First, a debriefing checklist was filled out to aid the pilots in an evaluation of the testing just completed. Their impressions were solicited relative to alerting format, attention-getting quality of the alerting tones, and possible masking effects caused by the concurrent introduction of voice alerts and ATC communications. After discussing these points to the satisfaction of the pilot and experimenter, an overflow logic evaluation was made by the pilot. For this portion of the debriefing, a combination of questionnaire and photographs were used to aid the pilot in evaluating three candidate systems for handling information overflow. Also elicited were the pilots' opinions on the application of color for coding priority level information.

The formal debriefing was generally followed by an informal discussion between pilot and experimenter where relevant pilot comments were recorded for further evaluation. Questions and comments regarding the purpose of the study, test methods, and quality of the simulation were also solicited.

5.5.3 TEST 5 – EVALUATION PROCEDURES

To make judgments on the alternate visual system concepts, the pilots were seated and shown photographs representing the different concepts to be evaluated (see Figures 5.2.5.1.1-1, -2 and 5.2.5.1.2-2). During this time the rationale and the operation of the visual system was explained; the relationship between the visual and aural systems was also described. Each option was explained in detail; the pilots were permitted to study each concept and to ask questions until satisfied that they understood its operation.

They were then given an evaluation questionnaire and asked to rate each concept. During the evaluation period the pilots could refer to the photographs and the test conductor to answer questions. They were told that the concepts were not necessarily complete and were asked to rate the concepts as they existed and to recommend changes which might improve them. The pilots' recommendations are summarized in the results section of this report and presented in detail in Appendix E. After completion of the questionnaire the responses were reviewed with the pilot to assure proper interpretation.

For the auditory evaluation, the pilots were seated between the speakers of a stereo tape player; they were advised they would be hearing alerting tones and would be asked some questions about them. They were told that one of the speakers would present the tones, the other, ATC background noise. The tones were presented one at a time for two seconds each and were separated by approximately six seconds; the sound level was set at approximately 75 db. The test conductor was present to answer questions and to make sure that the evaluation was properly completed.

5.6 PERFORMANCE MEASURES

The performance measures for the tests conducted in Phase I fall into two categories: those associated with the flight task and those associated with the alert systems. The parameters that reflect how well the pilot is performing the flight task were: pitch and roll error from the flight director, deviations from required speed, accuracy of responding to ATC queries the number of wheel and column reversals, and, time and accuracy in detecting targets in the external visual scene. The flight parameters were especially important for the time period immediately around an alert; each provided a measure of the efficiency and effectiveness of the pilot in performing the flight task. A second set of dependent variables, used to quantify the responses to the alerting system included: the time and accuracy of alert detection, the time and accuracy of the response to the alert, and the sequence in which the pilot performed the alert cancellation.

Finally, subjective data expressing the pilot's opinions about the various alerting systems were gathered for all test configurations. The pilots were asked to comment on and rate the effectiveness of the information display, clarity of the messages, format and color.

5.7 DATA REDUCTION AND ANALYSIS

Data resulting from the Phase I effort can be classified into two general categories: objective (or performance) data and subjective (questionnaire) data. A time-based tabulation of all events that occurred in the cab concerning switch and light states, displayed messages, and fault situation initiation was generated from the data. From the tabulation, sums, means and standard deviations were calculated for all the dependent variables. For tests 1 and 2 performance was investigated with respect to all the alerts and was also partitioned into the various alert categories; tests 3 and 4 did not partition the alerts; an analysis of variance was then be performed on the reduced data to determine if the various treatment conditions had a differential effect upon performance. The four objective tests was a factorial with repeated measures. The model and source table for this type of analysis is presented in Table 5.7-1.

5.7.1 TEST 1 – EXPERIMENTAL HYPOTHESES

The following are the hypotheses upon which test 1 was based: Each of the hypotheses can be equally stated for all three levels of alert, warning, caution and advisory.

1. There is no difference in detection time with and without an attention-getting device.

2. A master attenson in the 15^{o} field of view produces the same detection times as a flashing box on the information display.

3. Response times do not change with different attention-getting devices.

4. The attention-getting device has no effect on the distribution of missed alerts.

5. The type of attention-getting device used has no effect on flight performance.

Table 5.7-1. Sample of Variance Model and Summary Table for a Factorial Design With Repeated Measures (Example Is Two-Factor Experiment With Repeated Measures on Both Factors)

Model

$$X_{ijkm} = \mu + \alpha_i + \beta_j + \gamma_k + \alpha\beta_{ij} + \alpha\gamma_{ik} + \beta\gamma_{jk} + \alpha\beta\gamma_{ijk} + \epsilon_{ijkm}$$

Summary table

Source	Expected mean square	F ratio
A	$\sigma^2 + nb\,\sigma_{AS}^2 + nbs\,\sigma_A^2$	$F = MS_a/MS_{as}$
B	$\sigma^2 + na\,\sigma_{BS}^2 + nas\,\sigma_B^2$	$F = MS_b/MS_{bs}$
Subject(s)	$\sigma^2 + nab\,\sigma_S^2$	
AB	$\sigma^2 + n\,\sigma_{ABS}^2 + ns\,\sigma_{AB}^2$	$F = MS_{ab}/MS_{abs}$
AS	$\sigma^2 + nb\,\sigma_{AS}^2$	
BS	$\sigma^2 + na\,\sigma_{BS}^2$	
ABS	$\sigma^2 + n\,\sigma_{ABS}^2$	

6. There is no change in detection time when the location of the information display is moved from outside to inside the 15° field of vision.

7. Response time is the same for the two information display locations.

8. There is no difference in missed alerts between the two display locations.

9. Workload has no affect on detection time.

10. Workload has no effect on response time.

11. Error rate is not effected by workload.

12. Flight performance is not affected by workload.

13. Flight performance is not affected by the information display location.

14. Flight performance is not affected by the urgency of the alert.

5.7.2 TEST 2 — EXPERIMENTAL HYPOTHESES

The following are the hypotheses upon which test 2 was based: Each of the hypotheses can be equally stated for all three levels of alerts.

1. There is no difference in detection times when using the different master alerting formats.

2. There is no difference in response times when using the different master alerting formats.

3. The error distribution is not changed by changing master alerting formats.

4. The attention-getting device has no effect on the distribution of missed alerts.

5. The type of attention-getting device used has no effect on flight performance.

6. There is no change in detection time when the location of the information display is moved from outside to inside the 15° field of vision.

7. Response time is the same for the two information display locations.

8. There is no difference in missed alerts between the two display locations.

9. Workload has no effect on detection time.

10. Workload has no effect on response time.

11. Error rate is not affected by workload.

12. Flight performance is not affected by workload.

13. Flight performance is not affected by the information display location.

14. Flight performance is not affected by the urgency of the alert.

5.7.3 TEST 3 – EXPERIMENTAL HYPOTHESES

The following are the hypotheses upon which test 3 was based:

1. Performance on the ATC recognition and voice alert tasks would be substantially degraded when advisory communications are presented concurrently with voice alert messages.

2. When messages are presented concurrently with voice alerts the female controller voice would produce more confusion and masking effects due to the qualitative similarity of the competing speech sources. These effects would be manifested in the form of:

    a. longer alert response times

    b. more frequent errors in recognition of ATC messages.

3. The tone-voice alerting mode would be more effective as an attention-getting device than the voice-only alerting mode in terms of response time to failure annunciations.

4. Because of its attention-getting quality, the tone-voice alerting mode would cause more frequent errors on the ATC recognition task.

5. Any reductions in response time resulting from addition of redundant language would be lost due to the longer time required to annunciate the critical elements of the complete sentence messages.

6. Because of the longer duration of the complete sentence messages, the probability of confusion or masking between voice alerts and ATC messages would be greater than with the word/phrase messages. The reduced capacity of the pilot to discriminate between competing speech sources should result in longer response times and higher error frequencies on the ATC recognition task.

7. In attempting to identify ATC communications presented concurrently with voice alert messages, there were four basic response categories:

   a. correct readback,

   b. request message repeat,

   c. omission error,

   d. incorrect readback.

   It was hypothesized that the distribution of error types would vary across alerting modes and voice message formats. It was anticipated that the more serious types of error (omission and incorrect readback errors) would occur most frequently with the tone-voice mode and the complete sentence message format.

8. It was hypothesized that the word/phrase message format, preceded by an alerting tone would be most effective overall in terms of:

   a. response time to failure annunciations.

   b. performance on the flight task, i.e.:

      (1) tracking task accuracy.

      (2) recognition of ATC communications.

5.7.4 TEST 4 - EXPERIMENTAL HYPOTHESES

The following are the hypothesis upon which test 4 was based:

1. It was hypothesized that performance on the primary alert task and flight tasks would vary as a function of turbulence level. (Visual workload main effects).

2. It was hypothesized that performance on the primary alert task and the secondary ATC recognition task would vary as a function of alert - ATC message timing (auditory workload main effect).

3. It was hypothesized that the relative effectiveness of alerting modes would be dependent upon the level of visual and auditory task loading (alerting mode and workload interactions). No single alerting mode would be most effective under all combinations of environmental conditions.

   a. Under degraded auditory conditions (high auditory workload) it was hypothesized that the tone - visual alerting mode would:

      (1) minimize time required to identify and respond to failure annunciations.

      (2) minimize disruptive effects on performance of the ATC recognition task.

   b. Under conditions of High tracking task difficulty, it was hypothesized that modes relying on the auditory system as the primary source of information (voice only and tone - voice) would

      (1) minimize time required to identify and respond to failure annunciations.

      (2) minimize disruptive effects on performance of the visual task.

c. Under conditions of high auditory <u>and</u> high visual workload it was hypothesized that the combined tone - voice - visual operating mode would be most effective since it makes use of multiple sensory channels to provide fully or partially redundant information to the pilot.

5.8 TEST RESULTS

5.8.1 RESULTS OF ANALYSIS OF DATA FROM TEST 1

Since the work in this research program was exploratory in nature, an error probability of .10 was selected as a test for significance in the analysis of the data obtained. Some results reported as significant may be of insufficient magnitude to be of practical importance; this however may be a false assumption due to the nature of the tests. For example, when a pilot knows that alerts are going to occur, he has a tendency to respond more quickly than he would in a normal situation. Since the speed of response is bounded on the low side by physical parameters, smaller differences between treatments are to be expected. It is also expected that any differences found in this type of test would be magnified in real world situations.

5.8.1.1 DETECTION TIMES

The analysis of variance summary table for test one detection time is presented in Table 5.8.1.1-1. The main effect attributed to the attenson format was significant (F = 3.00 df 2,12). Using Duncan's New Multiple Range test it was discovered that the dual master was detected slower than the single master, but equal to the flashing master (2.81 sec vs. 2.45 sec vs. 2.59 sec). This effect is largely offset when looking at the significant interaction between attenson format and alert types (F = 5.28 df 4,24); this interaction shows that the differences in the detection times are contained mainly in the detection of the advisories. An illustration of this interaction is presented in Figure 5.8.1.1-1 and Table 5.8.1.1-2. The advisories are significantly different from the cautions and warnings. Also the advisory with the dual master has a mean detection time (4.19 sec) which is significantly slower than with the single master (2.97 sec). If the advisories are considered to be an alert with no master attenson, and are compared to those alerts with an attenson, the significant difference can be seen in Figure 5.8.1.1-2. This difference is also evident in the significant main effect attributed to the alert type (F = 71.87 df 2,12). The detection time for advisories (3.54 sec) was significantly longer than for either cautions (2.27 sec) or warnings (2.07 sec); the difference between cautions

Table 5.8.1.1-1. Test 1: Anova Summary Table for Detection Times

Source	Sum of squares	Degrees of freedom	Mean square	F ratio	Probability F exceeded
Mean	2650.83887	1	2650.83887	29.01505	0.00
Error	548.16504	6	91.36084		
Attenson format	9.24232	2	4.62116	3.00262	0.08*
Error	18.46849	12	1.53904		
Display format	16.19361	2	8.09681	1.40514	0.28
Error	69.14747	12	5.76229		
Attenson x display	3.15128	4	0.78782	0.19204	0.94
Error	98.45507	24	4.10229		
Workload	11.63372	1	11.63372	3.96510	0.09*
Error	17.60419	6	2.93403		
Attenson x workload	4.70389	2	2.35195	1.17643	0.34
Error	23.99069	12	1.99922		
Display x workload	0.84594	2	0.42297	0.39354	0.68
Error	12.89723	12	1.07477		
Attenson x display x workload	6.43328	4	1.60832	0.89320	0.48
Error	43.21524	24	1.80063		
Alert type	169.93672	2	84.96836	71.87662	0.00*
Error	14.18570	12	1.18214		
Attenson x alert type	26.24753	4	6.56188	5.28031	0.00*
Error	29.82500	24	1.24271		
Display x alert type	6.72503	4	1.68126	1.59658	0.20
Error	25.27285	24	1.05304		
Attenson x display x alert type	3.52166	8	0.44021	0.22490	0.98
Error	93.95464	48	1.95739		
Workload x alert type	1.03872	2	0.51936	0.24819	0.78
Error	25.11060	12	2.09255		
Attenson x workload x alert type	0.46651	4	0.11663	0.07768	0.98
Error	36.03158	24	1.50132		
Display x workload x alert type	0.90169	4	0.22542	0.15369	0.95
Error	35.20082	24	1.46670		
Attenson x display x workload x alert type	2.51196	8	0.31399	0.22947	0.98
Error	65.67969	48	1.36833		

*Significant at the 0.10 level or better.

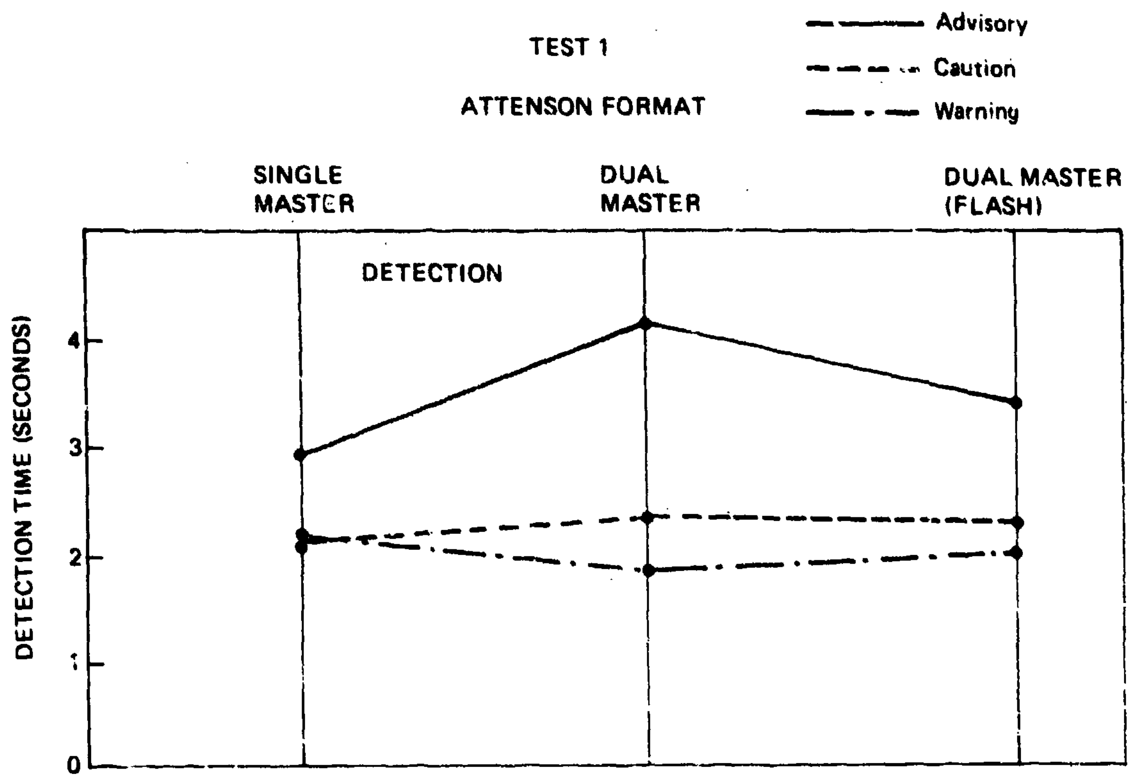


Figure 5.8.1.1-1. Detection Time as a Function of Attenson Format

Table 5.8.1.1-2. Mean Detection Times for Attenson Format by Alert-Type Interaction

Alert type	Attenson format		
	Single master (sec)	Dual master (sec)	Dual flashing (sec)
Advisory	2.97	4.19	3.44
Caution	2.15	2.35	2.31
Warning	2.23	1.92	2.06

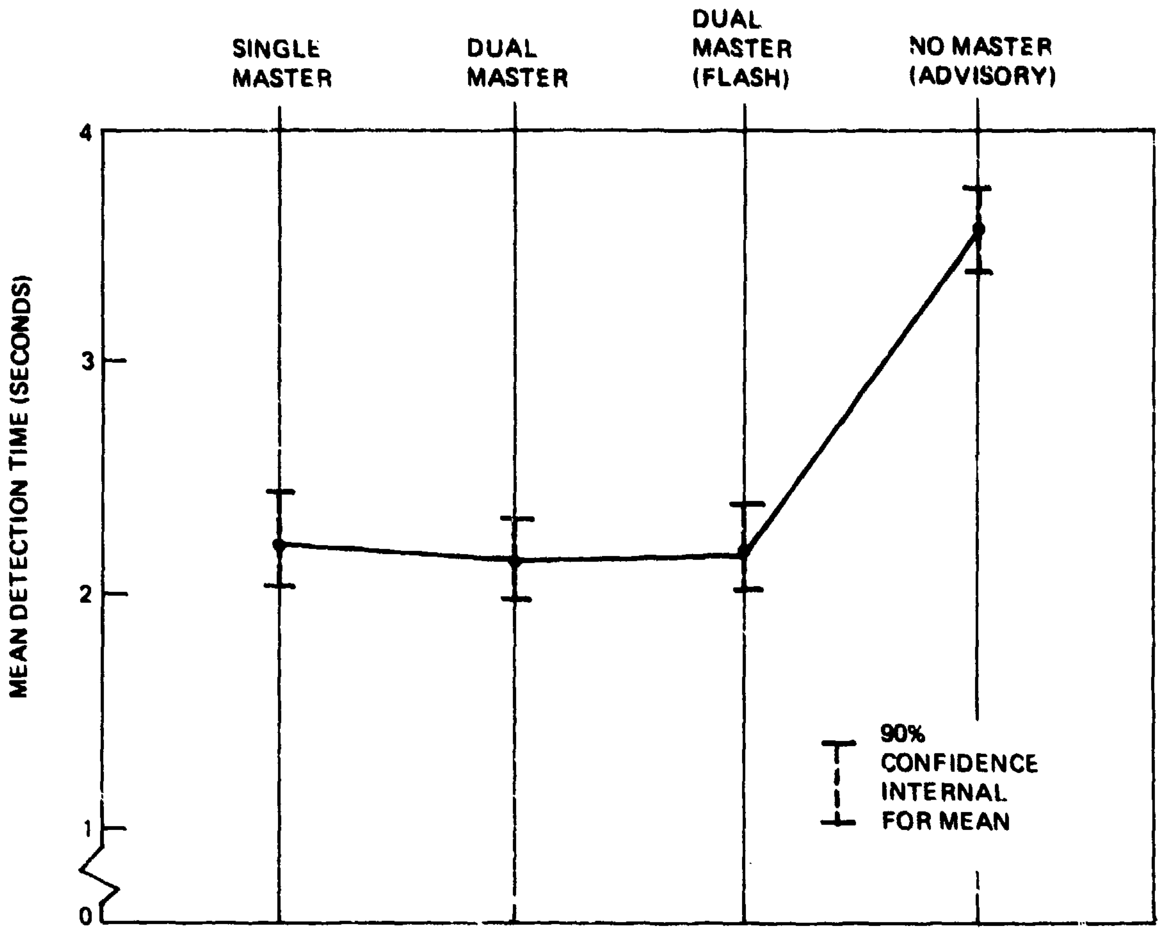


Figure 5.8.1.1-2. Mean Detection Time as a Function of Attenson Format

and warnings was not significant. Finally, the main effect attributed to workload was significant ($F = 3.97$ df 1,6); the detection time during high workload flights (2.80 sec) were significantly longer than on low workload flights (2.43 sec). No significant differences in mean detection times were observed as a function of display format (Figure 5.1.8.1-3).

5.8.1.2 RESPONSE TIMES

The analysis of variance summary table for test 1 response times is presented in Table 5.8.1.2-1. The main effect attributed to the alert type was significant ($F = 43.58$ df 2,12) as it was for detection times. However, unlike the results for detection time, a significant difference was detected among all three alert types. The mean response time to advisories (7.20 seconds) was significantly longer than the time for cautions (6.31 sec) which in turn was significantly longer than the time for warnings (5.23 sec). A significant interaction occurred between the attenson format and the target type ($F = 5.21$ df 4,24). Table 5.8.1.2-2 provides the cell means for this interaction; Figure 5.8.1.2-1 is a graphical representation. The mean response time for advisories under the dual attenson condition is responded to significantly slower than any other alert. The other two advisory (single master and dual master with flash) and all of the cautions were not significantly different in response time.

Finally the warnings were not significantly different among themselves but were responded to significantly faster than any of the other alerting conditions. A second interaction that proved to be significant was the interaction between display format and alert types ($F = 3.04$ df 4,24). Table 5.8.1.2-3 presents the cell means for this interaction; Figure 5.8.1.2-2 provides the graphic representations. The major significant difference illustrated by this interaction is that when there is an attenson present, the fastest response time is achieved by having the messages appear at the top of the display. These conditions resulted in mean response times which were not significantly different from each other, but were significantly different from the rest of the treatment means.

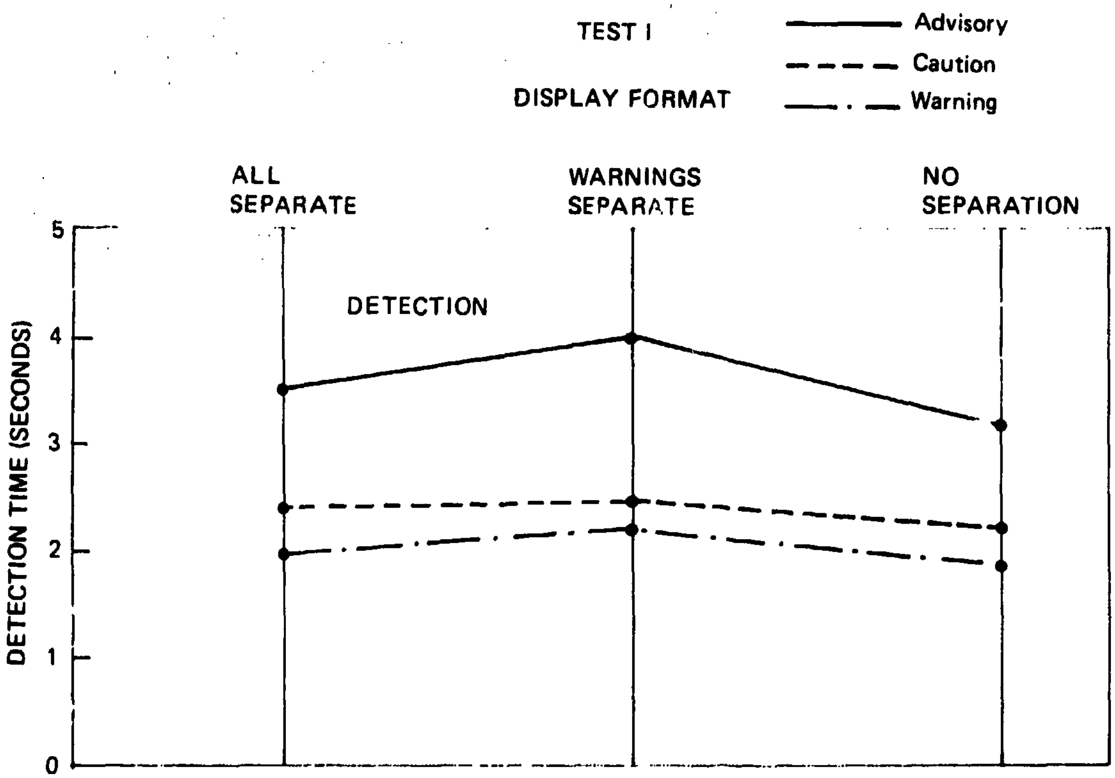


Figure 5.8.1.1-3. Detection Time as a Function of Display Format

Table 5.8.1.2-1. Anova Summary Table for Test 1 Response Times

Source	Sum of squares	Degrees of freedom	Mean square	F ratio	Probability F exceeded
Mean	14728.88672	1	14728.88672	203.21625	0.00
Error	434.87329	6	72.47888		
Attenson format	31.97916	2	15.98958	0.89930	0.43
Error	213.36089	12			
Display format	54.73848	2	27.36924	2.05302	0.17
Error	159.97444	12	13.33120		
Attenson x display	18.72372	4	4.68093	0.35990	0.83
Error	312.14975	24	13.00624		
Workload	16.91217	1	16.91217	4.31991	0.08*
Error	23.48959	6	3.91493		
Attenson x workload	10.05143	2	5.02572	1.35105	0.29
Error	44.63834	12	3.71986		
Display x workload	0.02121	2	0.01060	0.00326	0.99
Error	39.06429	12	3.25536		
Attenson x display x workload	12.37762	4	3.09440	0.83300	0.51
Error	89.15465	24	3.71478		
Alert type	240.17386	2	120.08693	43.58478	0.00*
Error	33.06299	12	2.75525		
Attenson x alert type	57.92708	4	14.48177	5.21322	0.00*
Error	66.66949	24	2.77790		
Display x alert type	27.13592	4	6.78398	3.03963	0.03*
Error	53.56424	24	2.23184		
Attenson x dislay x alert type	5.90974	8	0.73872	0.36294	0.09
Error	97.69718	48	2.03536		
Workload x alert type	2.71825	2	1.35913	0.65334	0.53
Error	24.96317	12	2.08026		
Attenson x workload x alert type	3.59728	4	0.89932	0.34788	0.84
Error	62.04324	24	2.58513		
Display x workload x alert type	4.82396	4	1.20599	0.52278	0.72
Error	55.36458	24	2.30686		
Attenson x display x workload x alert type	11.40337	8	1.42542	0.52278	0.85
Error	138.72610	48	2.89013		

*Significant at 0.10 level or better.

Table 5.8.1.2-2. Mean Response Times for Attenson Format by Alert-Type Interaction

Alert type	Attenson format			
	Single (sec)	Dual (sec)	Dual/flash (sec)	
Advisory	A 6.45	B 8.28	C 6.82	
Caution	D 6.03	E 6.56	F 6.33	
Warning	G 5.24	H 4.99	I 5.44	

B is greater than all others.

A, C, D, E, and F are not different and are significantly greater than the rest.

G, H, and I are not different.

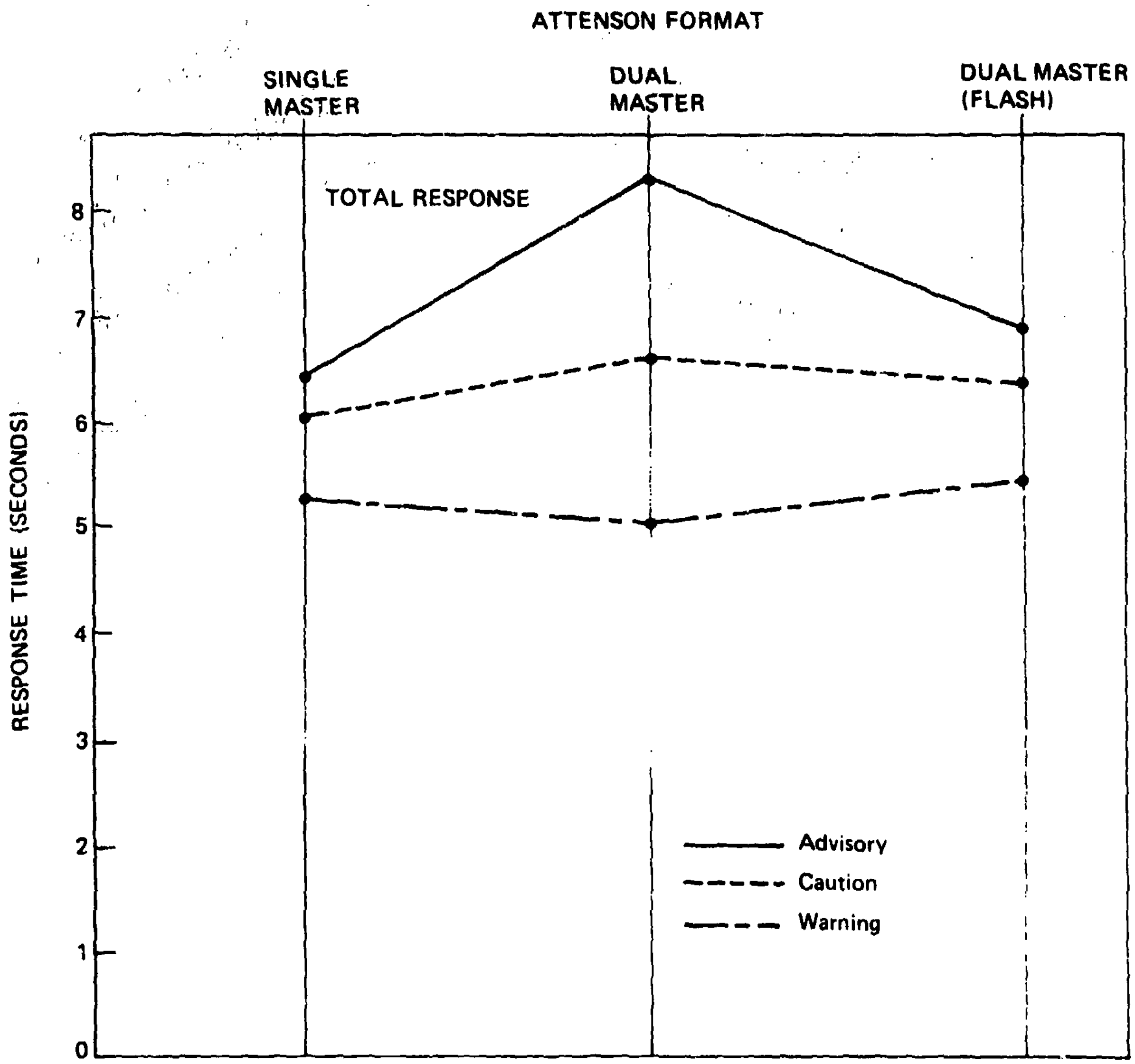


Figure 5.8.1.2-1. Response Time as a Function of Attenson Format

Table 5.8.1.2-3. Mean Response Times for Display Format by Alert-Type Interaction

Alert type	Display format		
	All categories separate (sec)	Warnings separate (sec)	No separation (sec)
Advisory	6.87	7.94	6.78
Caution	6.75	6.69	5.48
Warning	5.25	5.42	5.01

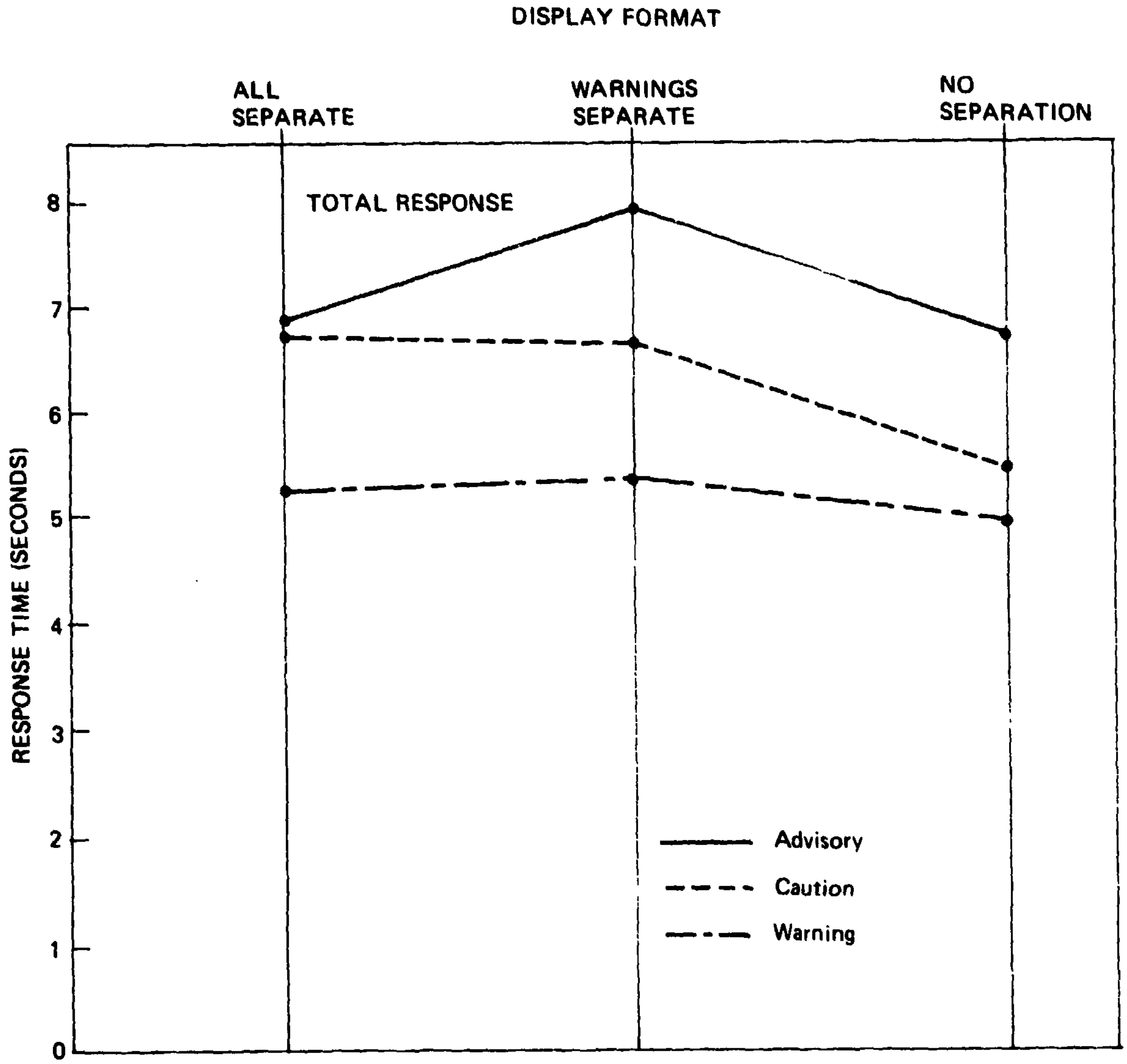


Figure 5.8.1.2-2. Response Time as a Function of Display Format

The main effect attributed to workload again was significant (F = 4.32 df 1,6). The mean response time during the high workload flights (6.45 sec) was significantly greater than the mean time for low workload flights (6.01 sec).

5.8.1.3 MISSED ALERTS

In an experiment of this type where the pilot is primed to detect alerts, it is very difficult to produce conditions which will result in a large number of missed alerts. For test one there were a total of 1512 possible alerts of which only 30 were missed, (2% of the total). The most significant contributor to the number of missed alerts was the presence/absence of the attenson. Advisories which did not have an attenson were missed 4.8 percent of the time (24 out of 504). Cautions were missed 1 percent (5 out of 504) and warnings only .2 percent 1 out of 504); this difference, which can be seen in Figure 5.1.8.3-1, was highly significant ($X^2 = 30.2$ $p < .001$). The other test variables did not have a significant effect on missed alerts; figures 5.1.8.3-2, -3, -4 present the distributions associated with these variables.

5.8.1.4 PILOT PREFERENCES

The questionnaire given to the pilots at the end of test 1 is contained in this report in Appendix B. The pilots showed no clear preference for any of the attenson formats. They rated the flashing attenson slightly higher in attention-getting quality (mean score 9.1) than the single alert (8.0), but all three ratings were acceptable. Four of the pilots commented that although the flash was more attention-getting, it may be too distracting; this comment is reflected in the ratings on disruption. The average rating for the single and dual alerts were 8.0 and 7.7 respectively, however the flashing attenson was rated at 5.1. The location of the alerts was considered appropriate with an average rating of 8.3 for all three concepts.

Although the display format which entered all new messages at the top of the display (chronologically ordered) was rated very high (8.3) as an aid in finding the most recent alert as compared to messages going into their own catagory (4.9) or warnings in their own catagory (5.7), the order of preference was reversed when rated for the ability to assess aircraft status.

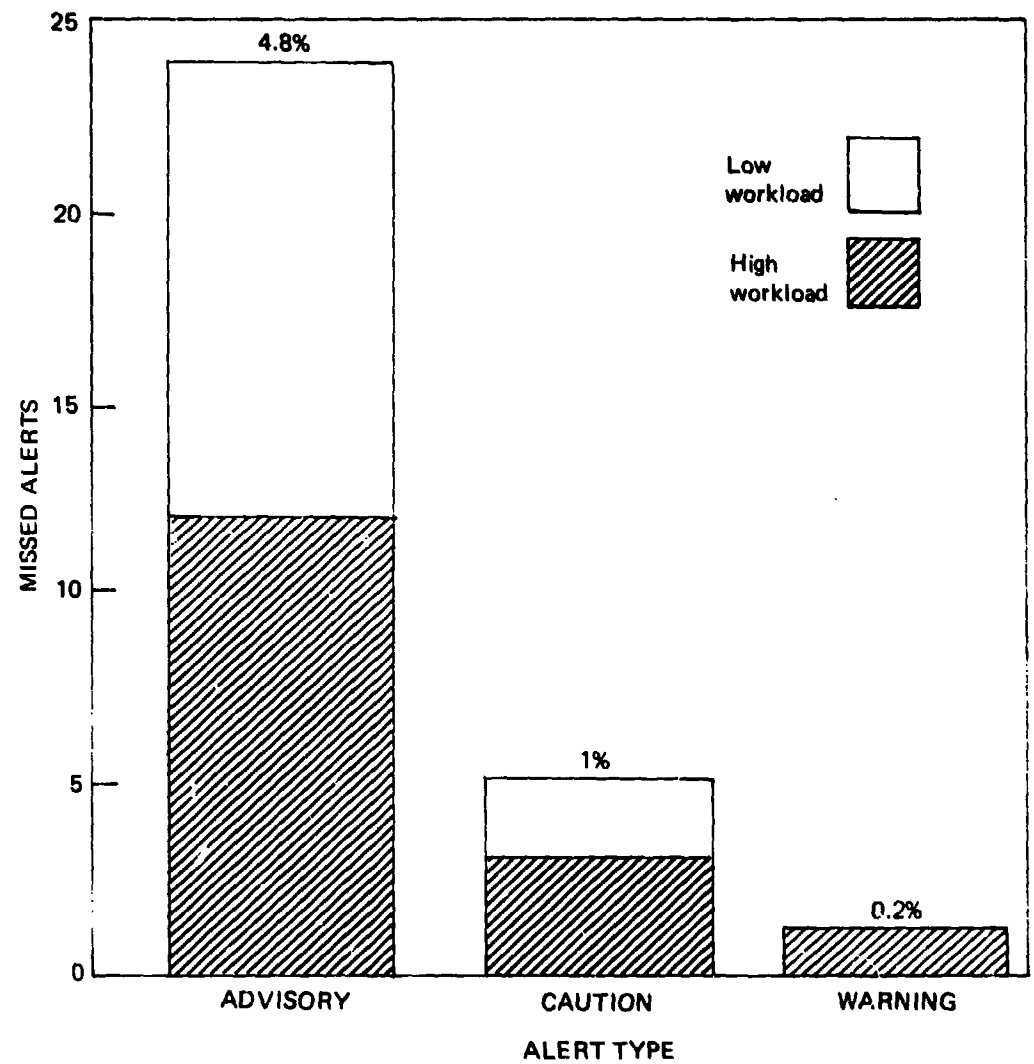


Figure 5.8.1.3-1. Missed Alerts as a Function of Alert Type

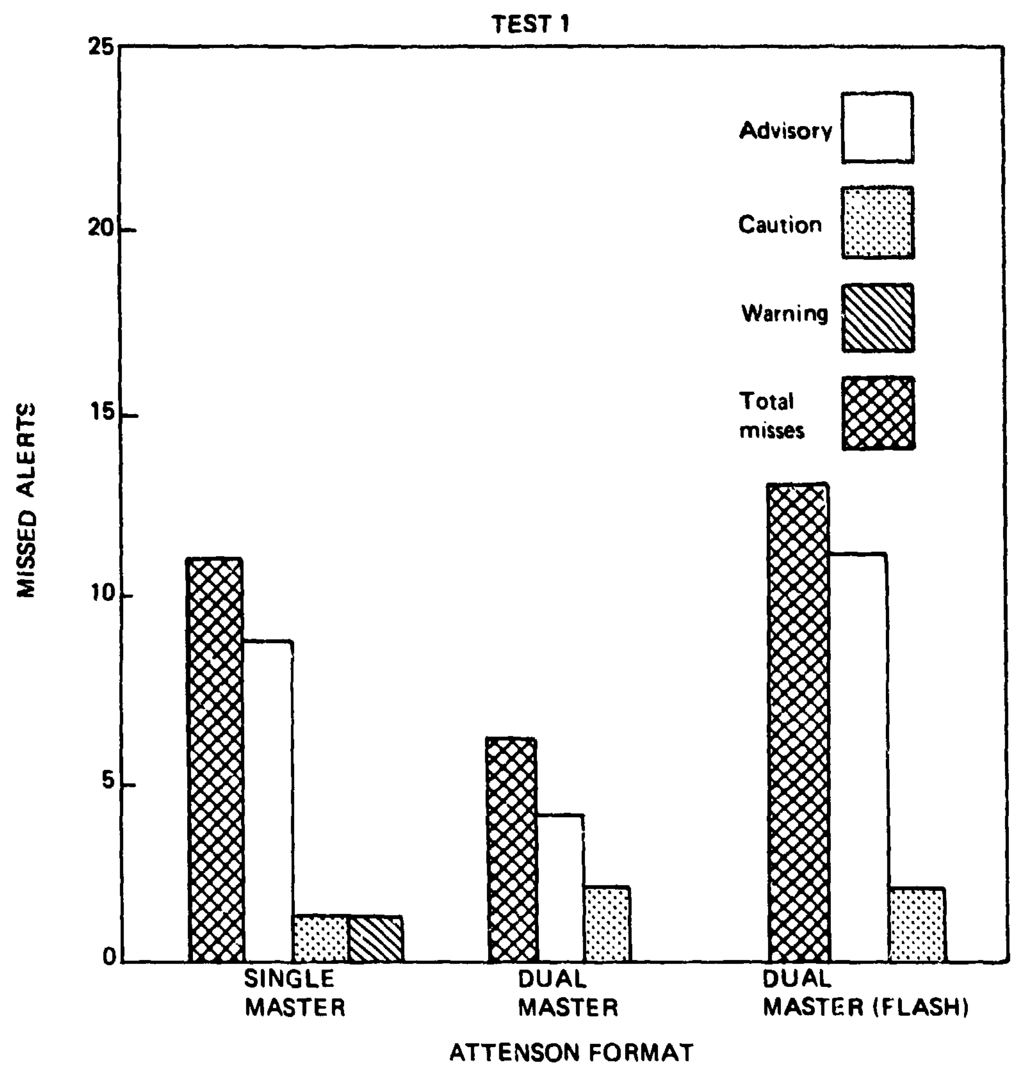


Figure 5.8.1.3-2. Missed Alerts as a Function of Attenson Format

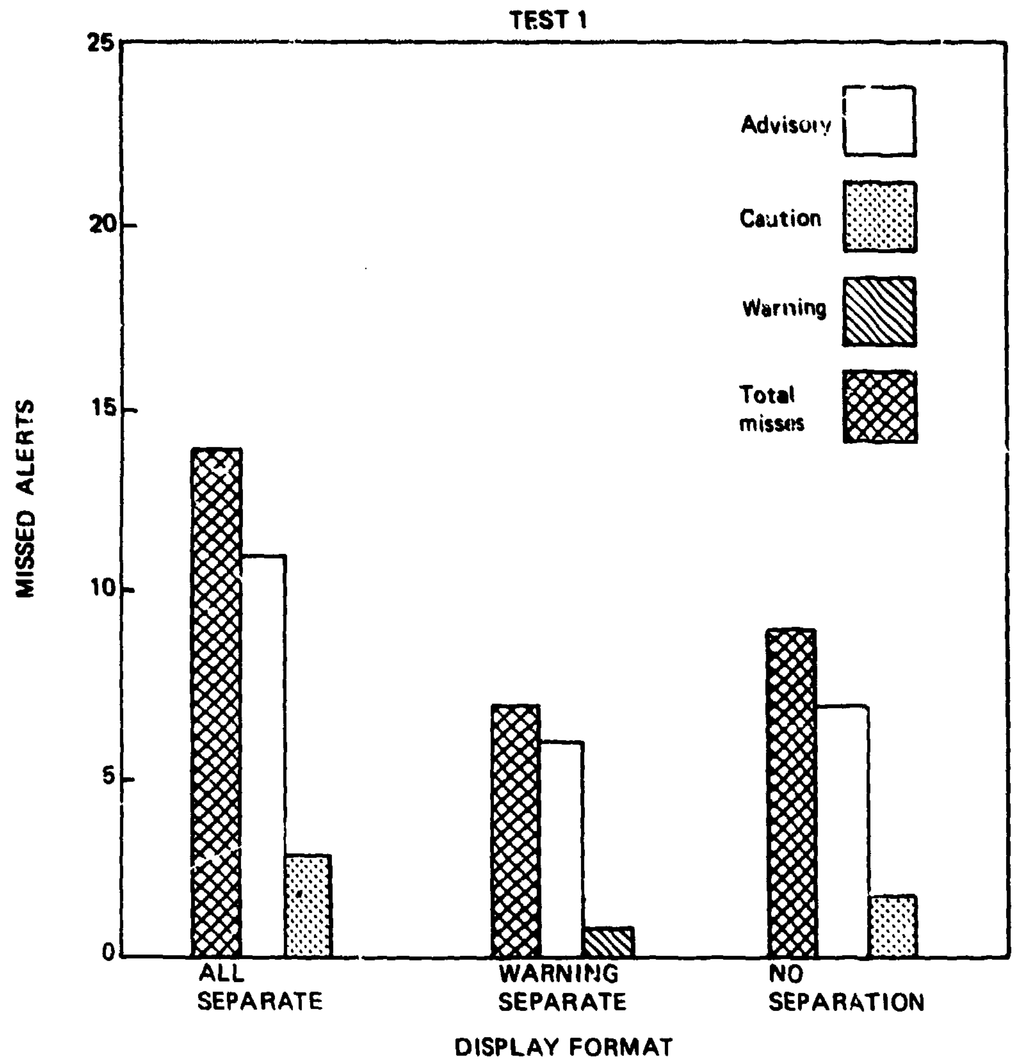


Figure 5.8.1.3-3. Missed Alerts as a Function of Display Format

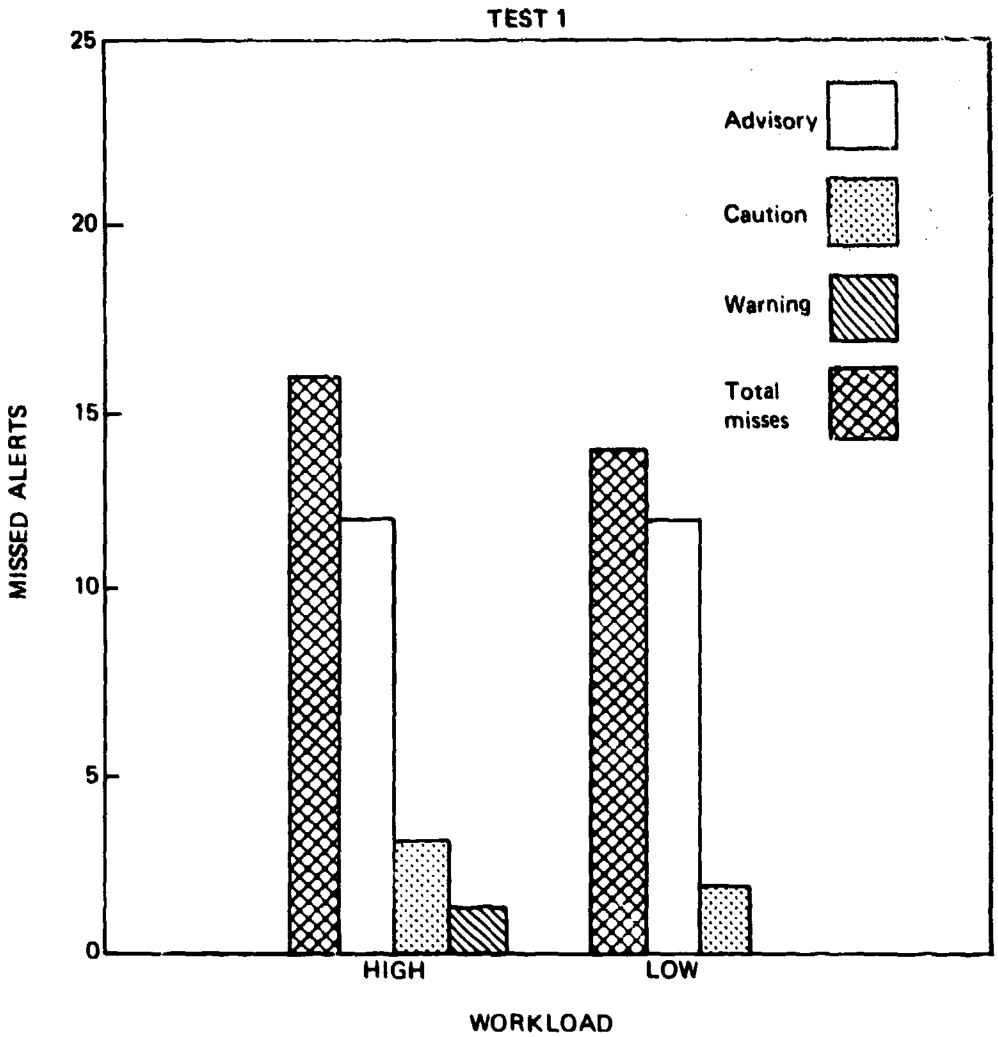


Figure 5.8.1.3-4. Missed Alerts as a Function of Workload

The pilots preferred categorization (7.7) to chronological ordering (6.2) and felt that they could assess aircraft status better with message catagorization (6.9 vs 5.7).

When reporting the five features they liked best, four of the seven pilots preferred the three color system. The detail of information was mentioned twice; all other identified features were mentioned only once.

Preferred changes that were mentioned included the addition of a cancel/recall function (4 pilots), and the addition of a blue master attenson for advisories (2 pilots).

As for features the pilots disliked, none received more than a single mention.

5.8.2 RESULTS OF DATA ANALYSIS FROM TEST 2

5.8.2.1 DETECTION TIMES

The analysis of variance for test two detection times is summarized in Table 5.8.2.1-1. Alerts were detected significantly faster (F = 11.70 df 1,5) when there was a master attenson (1.95 sec) than when there was none (2.55 sec). Alerts were also detected significantly faster (F = 10.62 df 1,5) when there was a flashing box (2.19 sec) than when there was none (2.32 sec). Since this may be misleading because the master attenson was also present for half of the trials, the interaction between the attenson and flashing box was also investigated. In fact there was a significant interaction (F = 4.48 df 1,5) which is illustrated in Figure 5.8.2.1-1. The test condition with neither a master attenson or a flashing box was significantly slower (2.68 sec) than all the other conditions. When the flashing box was added, the detection time improved (2.42 sec) but was still significantly slower than either of the conditions with a master attenson. There was no difference with the master attenson as the only attention getter (1.96 sec) or when both the master attenson and flashing box were used (1.94 sec). The main effect attributed to workload was significant (F = 79.37 df 1,5) with high workload resulting in a longer detection time (2.38 sec) than low workload (2.13 sec). The alert type also showed a significant main effect for detection time (F = 10.77 df 2,10). As in test one, the advisories were detected significantly slower (2.85 sec) than either cautions (2.08 sec) or warnings (1.83 sec); detection time for cautions and warnings were not significantly different. The significant interactions between the master visual attenson and the type of alert (F = 16.82 df 2,10) again illustrates the value of the master light. Although the treatment means consisted of trials with and without the flashing box, a significant advantage can be seen when the master visual light was used (Figure 5.8.2.1-2). The mean detection time for both cautions and warnings were significantly faster with the master attenson than for the other treatment means.

There were some higher order interactions that were also significant; however, because of the complexity existing in these interactions, they were extremely difficult to interpret and use. Since the master attenson and the flashing

Table 5.8.2.1-1. Anova Summary Table for Test 2 Detection Times

Source	Sum of squares	Degrees of freedom	Mean square	F ratio	Probability F exceeded
Mean	2225.11304	1	2225.11304	113.02164	0.00
Error	98.43748	5	19.68750		
Location	0.48727	2	0.24363	0.09715	0.90
Error	25.07861	10	2.50786		
Attenson format	35.29986	1	35.29986	11.69597	0.01*
Error	15.09061	5	3.01812		
Location x attenson	5.79630	2	2.89815	0.72587	0.50
Error	39.92661	10	3.99266		
Box	1.56133	1	1.56133	10.61804	0.02*
Error	0.73522	5	0.14704		
Location x box	10.28124	2	5.14062	2.40593	0.14
Error	21.36641	10	2.13664		
Attenson x box	1.90709	1	1.90709	4.47864	0.08*
Error	2.12909	5	0.42582		
Location x attenson x box	1.22270	2	0.61135	0.19283	0.82
Error	31.70367	10	3.17037		
Workload	7.43269	1	7.43269	79.37555	0.00*
Error	0.46820	5	0.09364		
Location x workload	3.96358	2	1.98179	0.66501	0.53
Error	29.80071	10	2.98007		
Attenson x workload	0.00469	1	0.00469	0.00131	0.97
Error	17.87870	5	3.57574		
Location x attenson x workload	9.36952	2	4.68476	2.42477	0.13
Error	19.32044	10	1.93204		
Box x workload	2.05523	1	2.05523	1.38767	0.29
Error	7.40532	5	1.48106		
Location x box x workload	1.02613	2	0.51306	0.47880	0.63
Error	10.71559	10	1.07155		
Attenson x box x workload	7.10838	1	7.10838	3.00292	0.14
Error	11.83580	5	2.36716		

*Significant at 0.10 level or better

Table 5.8.2.1-1. Anova Summary Table for Test 2 Detection Times (Concluded)

Source	Sum of squares	Degrees of freedom	Mean square	F ratio	Probability F exceeded
Location x attenson x box x workload	0.47416	2	0.23708	0.28093	0.76
Error	8.43918	10	0.84392		
Alert type	81.93771	2	40.96866	10.77352	0.00*
Error	38,02737	10	3.80274		
Location x alert type	14.56619	4	3.64155	2.41680	0.08*
Error	30.13520	20	1.50676		
Attenson x alert type	33.63356	2	16.81678	8.45489	0.00*
Error	19,89000	10	1.98900		
Location x attenson x alert type	11.25547	4	2.81387	0.79393	0.54
Error	70.88480	20	3.54424		
Box x alert type	8.87672	2	4.43836	3.47838	0.07*
Error	12.75984	10	1.27598		
Location x box x alert type	40.07364	4	10.01846	4.76299	0.00
Error	42.06791	20	2.10340		
Attenson x box x alert type	1.35113	2	0.67557	0.66787	0.53
Error	10.11519	10	1.01152		
Location x attenson x box x alert type	17.68275	4	4.42069	2.91096	0.04*
Error	30.37276	20	1.51864		
Workload x alert type	0.27078	2	0.13530	0.12216	0.88
Error	11.08270	10	1.10827		
Location x workload x alert type	1.63228	4	0.40807	0.13263	0.96
Error	61.53366	20	3.07668		
Attenson x workload x alert type	7.02684	2	3.51342	1.98079	0.18
Error	17.73750	10	1.77375		
Location x attenson x workload x alert type	28.54462	4	7.13615	2.62238	0.06*
Error	54.42505	20	2.72125		
Box x workload x alert type	4.15355	2	2.07677	1.05133	0.38
Error	19.75373	10	1.97537		
Location x box x workload x alert type	9.08467	4	2.27117	2.04963	0.12
Error	22.16175	20	1.10809		
Attenson x box x workload x alert type	7.76877	2	3.88439	2.81329	0.10*
Error	13.80729	10	1.38073		
Location x attenson x box x workload x alert type	4.22060	4	1.05515	0.67011	0.62
Error	31.49196	20	1.57460		

*Significant at 0.10 level or better

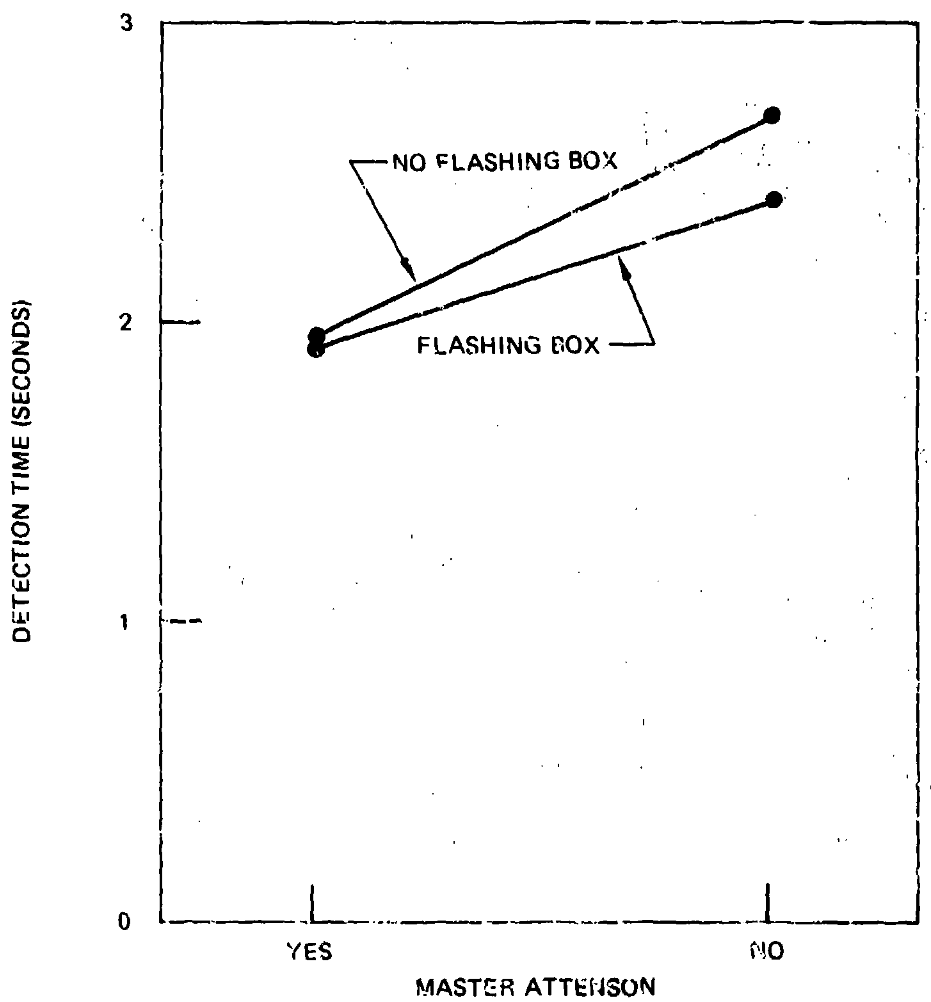


Figure 5.8.2.1-1. Detection Time Interaction Between Master Attensor and Flashing Box

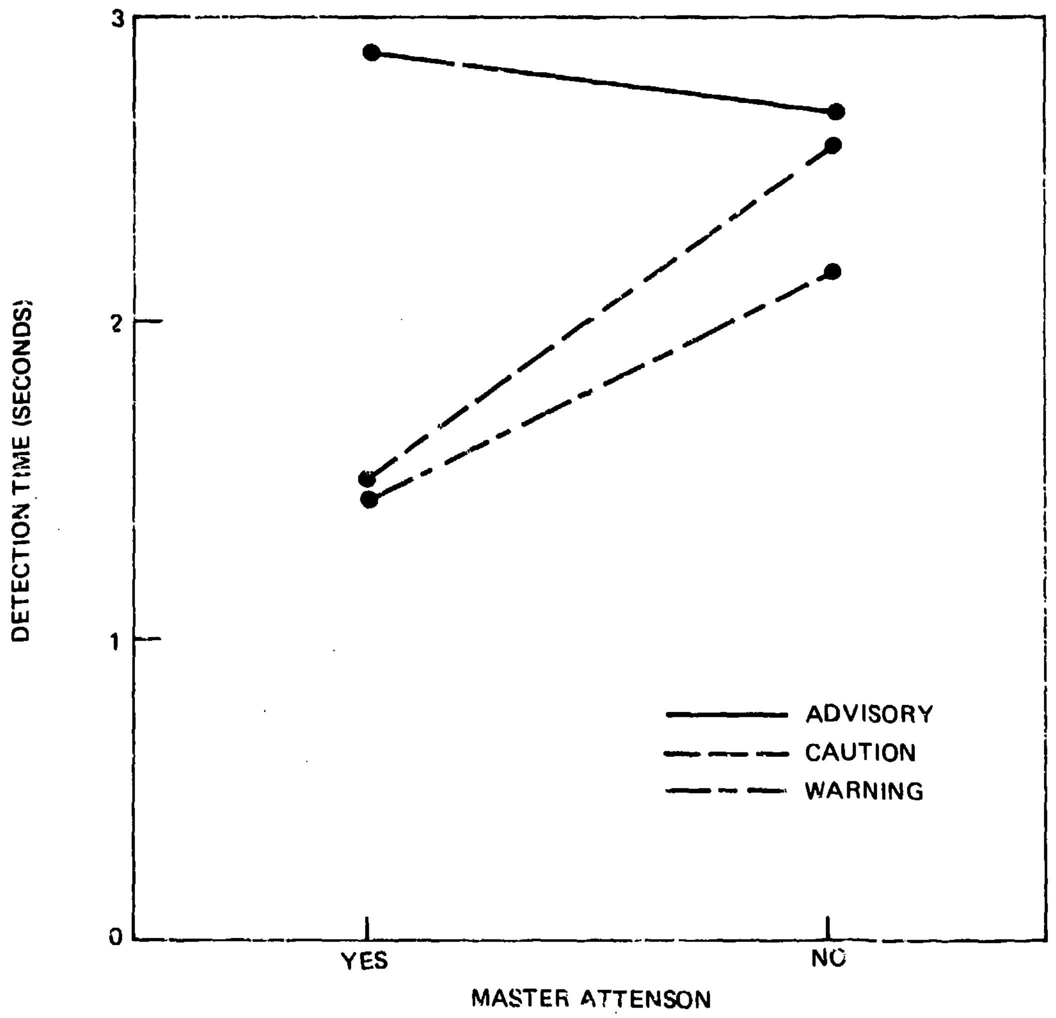


Figure 5.8.2.1-2. Detection Time as Function of Master Attenson and Target Type Across All Other Variables in Test 2

box were both used to get the pilots attention, it was possible to compare the detection times of both. The significant ($F = 4.42$ df 4,20) interaction between display location, master visual attenson, flashing box and alert type is presented in Figures 5.8.2.1-3, -4 and -5 respectively as an example of a higher order interaction combining these attention-getters. One finding supported by this interaction is the value of the master visual attenson. The mean detection time for both cautions and warnings with a master were not different from each other; however, when the flashing box was included they were significantly different from all other treatments. The mean response times for the conditions using the flashing box alone were not as consistent. In fact, when the cautions or warnings were on the center display the mean detection times were significantly slower than any of the conditions using a master attenson (3.45 sec and 3.05 sec) with the display in front of the pilot, the flashing box was found to be as attention-getting as the master visual attenson.

This interaction demonstrates a difficulty with this type of experiment. Figure 5.8.2.1-5 shows that the mean detection time for the advisories were highly dependent on the treatment conditions. Detection times were significantly longer for two of the four conditions. A possible explanation for this is that the treatment conditions themselves affected the pilots scan pattern in a systematic manner. With the major alerts (warnings and cautions) separated from the advisories the pilot scanned the advisory display less often and their detection times became longer. Even though this type of data may be an artifact of the experimental conditions, it illustrates what can happen in real world systems if the human element is neglected in the design.

5.8.2.2 RESPONSE TIME

Table 5.8.2.2-1 summarizes the analysis of variance for test two response times. The main effect attributed to alert type was significant ($F = 14.69$ df 2,10) as it was for detection times; however, the emphasis in responding differed from that of detecting the alerts.

Advisories were detected significantly slower than warnings or cautions; warnings had a mean response time significantly faster (5.05 sec) than

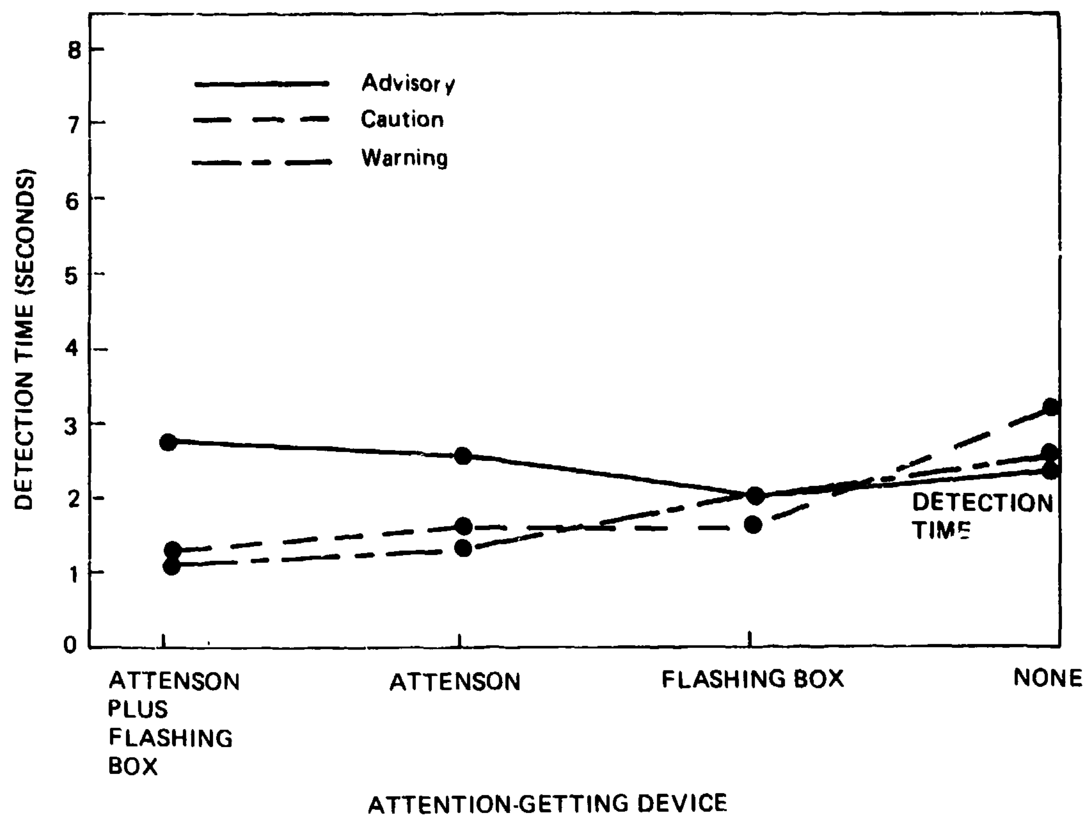


Figure 5.8.2.1-3. Test 2: Detection Time as Function of Attention-Getting Device All Alerts on Center Display

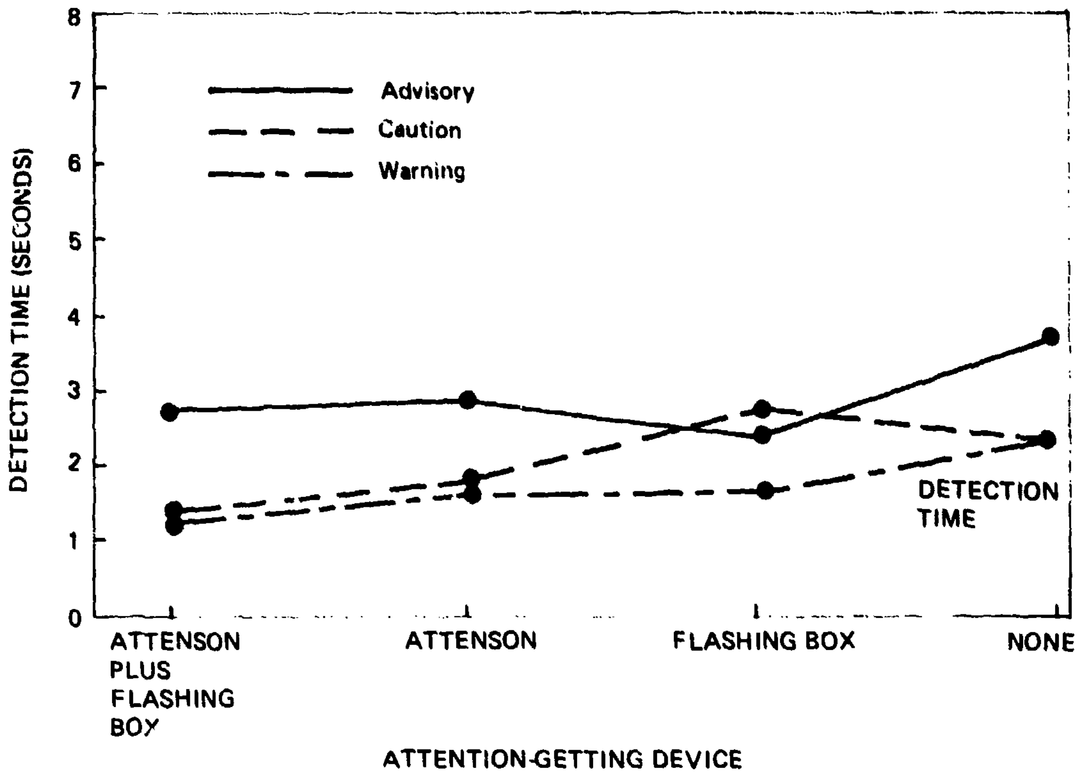


Figure 5.8.2.1-4. Test 2: Detection Time as Function of Attention-Getting Device—Warnings on Pilot Display

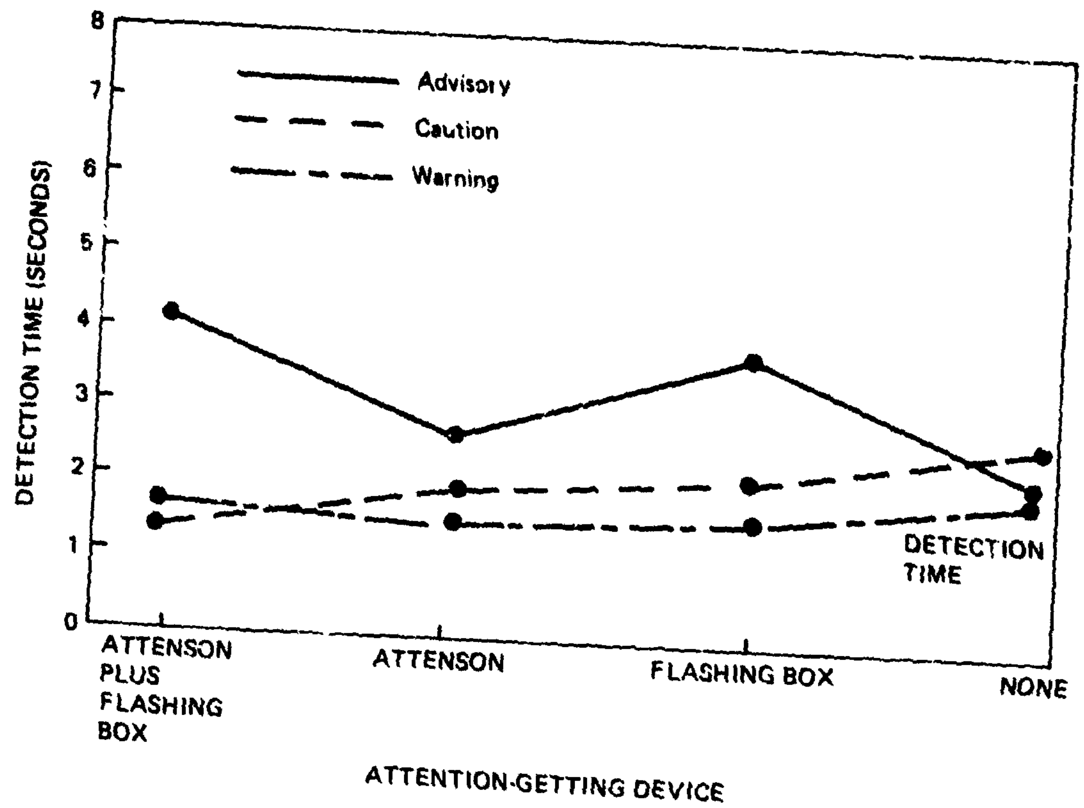


Figure 5.8.2.1-5. Test 2: Detection Time as Function of Attention-Getting Device—Warning and Cautions on Pilot Display

Table 5.8.2.2-1. Anova Summary Table for Test 2 Response Times

Source	Sum of squares	Degrees of freedom	Mean square	F ratio	Probability F exceeded
Mean	15727.64453	1	15727.64453	350.05212	0.00
Error	224.64717	5	44.92944		
Location	17.18568	2	8.59284	0.81999	0.46
Error	104.79192	10	10.47919		
Attenson format	2.28115	1	2.28115	0.30780	0.60
Error	37.05596	5	7.41119		
Location x attenson	12.48304	2	6.24512	0.97871	0.40
Error	63.77308	10	6.37731		
Box	2.97572	1	2.97572	0.42860	0.54
Error	34.71444	5	6.94289		
Location x box	1.58067	2	0.79033	0.07286	0.93
Error	108.46937	10	10.84694		
Attenson x box	0.95485	1	0.95484	0.55493	0.49
Error	8.60338	5	1.72068		
Location x attenson x box	41.33916	2	20.66958	2.89047	0.10*
Error	71.50952	10	7.15095		
Workload	32.45915	1	32.45915	5.29885	0.07*
Error	30.62846	5	6.12569		
Location x workload	40.11121	2	20.05560	5.75767	0.02*
Error	34.83282	10	3.48328		
Attenson x workload	0.53116	1	0.53116	0.37374	0.56
Error	7.10598	5	1.42120		
Location x attenson x workload	34.59525	2	17.29762	4.17753	0.04*
Error	41.40631	10	4.14063		
Box x workload	4.78845	1	4.78845	2.05696	0.21
Error	11.63961	5	2.32792		
Location x box x workload	10.41762	2	5.20881	2.80621	0.10*
Error	18.56173	10	1.85617		
Attenson x box x workload	2.24035	1	2.24035	0.43957	0.53
Error	25.48371	5	5.09674		

*Significant at 0.10 level or better.

Table 5.8.2.2-1. Anova Summary Table for Test 2 Response Times (Concluded)

Source	Sum of squares	Degrees of freedom	Mean square	F ratio	Probability F exceeded
Location x attenson x box x workload	1.16792	2	0.58396	0.37663	0.49
Error	15.50500	10	1.55050		
Alert type	208.75732	2	104.37866	14.69378	0.00*
Error	71.03595	10	7.10359		
Location x alert type	19.72819	4	4.93205	2.37472	0.08*
Error	41.53793	20	2.07690		
Attenson x alert type	19.03846	2	9.51923	3.05572	0.09*
Error	31.15218	10	3.11522		
Location x attenson x alert type	25.04624	4	6.26156	1.32061	0.29
Error	94.82832	20	4.74142		
Box x alert type	21.72777	2	10.86389	3.57691	0.06*
Error	30.37224	10	3.0372		
Location x box x alert type	41.43655	4	10.35914	3.50747	0.02*
Error	59.06902	20	2.95345		
Attenson x box x alert type	16.77257	2	8.38629	5.86847	0.02*
Error	14.29042	10	1.42904		
Location x attenson x box x alert type	27.95685	4	6.98921	2.42120	0.06*
Error	57.73345	20	2.88667		
Workload x alert type	1.11862	2	0.55931	0.17298	0.84
Error	32.33472	10	3.23347		
Location x workload x alert type	5.09377	4	1.27344	0.27142	0.89
Error	93.83466	20	4.69173		
Attenson x workload x alert type	10.49886	2	5.24943	4.54979	0.03*
Error	11.53775	10	1.15377		
Location x attenson x workload x alert type	28.27876	4	7.06969	1.80741	0.16
Error	78.23005	20	3.91150		
Box x workload x alert type	21.78006	2	10.89003	5.67647	0.02*
Error	19.18451	10	1.91845		
Location x box x workload x alert type	13.27473	4	3.31868	2.99837	0.04*
Error	22.13654	20	1.10683		
Attenson x box x workload x alert type	18.56601	2	9.28300	3.81618	0.05*
Error	24.32536	10	2.43254		
Location x attenson x box x workload x alert type	4.41753	4	1.10438	0.43405	0.78
Error	50.88723	20	2.54436		

*Significant at 0.10 level or better

cautions (6.35 sec). The mean response times for cautions and advisories were not significantly different. Pilot workload had a significant effect on the response time (F = 5.30 df 1,5); mean response times for high workload flights were slower (6.28 sec) than those for low workload flights (5.74 sec).

A number of the interactions containing workload as one of the variables were significant. Care should be used when interpreting these results, especially when display location is also one of the variables, because the lack of an attenson combined with the center display location created a very difficult detection and response task. An example of this effect can be seen in the significant interaction between pilot workload and display location (F = 5.76 df 2,10). Examination of the mean response times in Table 5.8.2.2-2 and Figure 5.8.2.2-1 reveals that the only significantly different treatment condition was the high workload flights when all alerts were presented on the central display (6.91 sec). A deeper look at the data revealed that the major contributers to this mean were: the mean response time to advisories when there was a master attenson for the cautions and warnings (8.84 sec); and the mean response times to the cautions and warnings when there was no master attenson (7.53 sec and 7.11 sec). The other mean response times were comparable to the rest of the treatment conditions. The master attenson and the flashing box can be combined for comparison since both were used as attention-getters with the latter used to aid the pilot in finding the most recent alert.

The significant interaction (F = 2.42 df 4,20) between display location, attention-getter and alert type is presented in Figures 5.8.2.2-2, -3 and -4 respectively. This higher order interaction was also difficult to interpret completely, however there are some important points that should be noted. Response to warnings is consistent when there is a master visual attenson (4.67 sec to 5.07 sec). The location of the display did not have a significant effect. With no attenson the flashing box was an adequate attention-getter as long as the warning was presented on a central display in front of the pilot. However, mean responses to warnings on the central display with no attenson were significantly slower (6.27 sec with a flashing box and 6.57 sec with no box). Mean response times to cautions exhibited the same characteristics as in the first experiment; with the attenson alone or

Table 5.8.2.2-2. Mean Response Times in Seconds for Interaction Between Pilot Workload and Display Location

Display location	Pilot workload	
	High	Low
All alerts in center	6.91	5.62
Warnings on pilot display	5.79	5.95
Warnings and cautions on pilot display	6.16	5.66

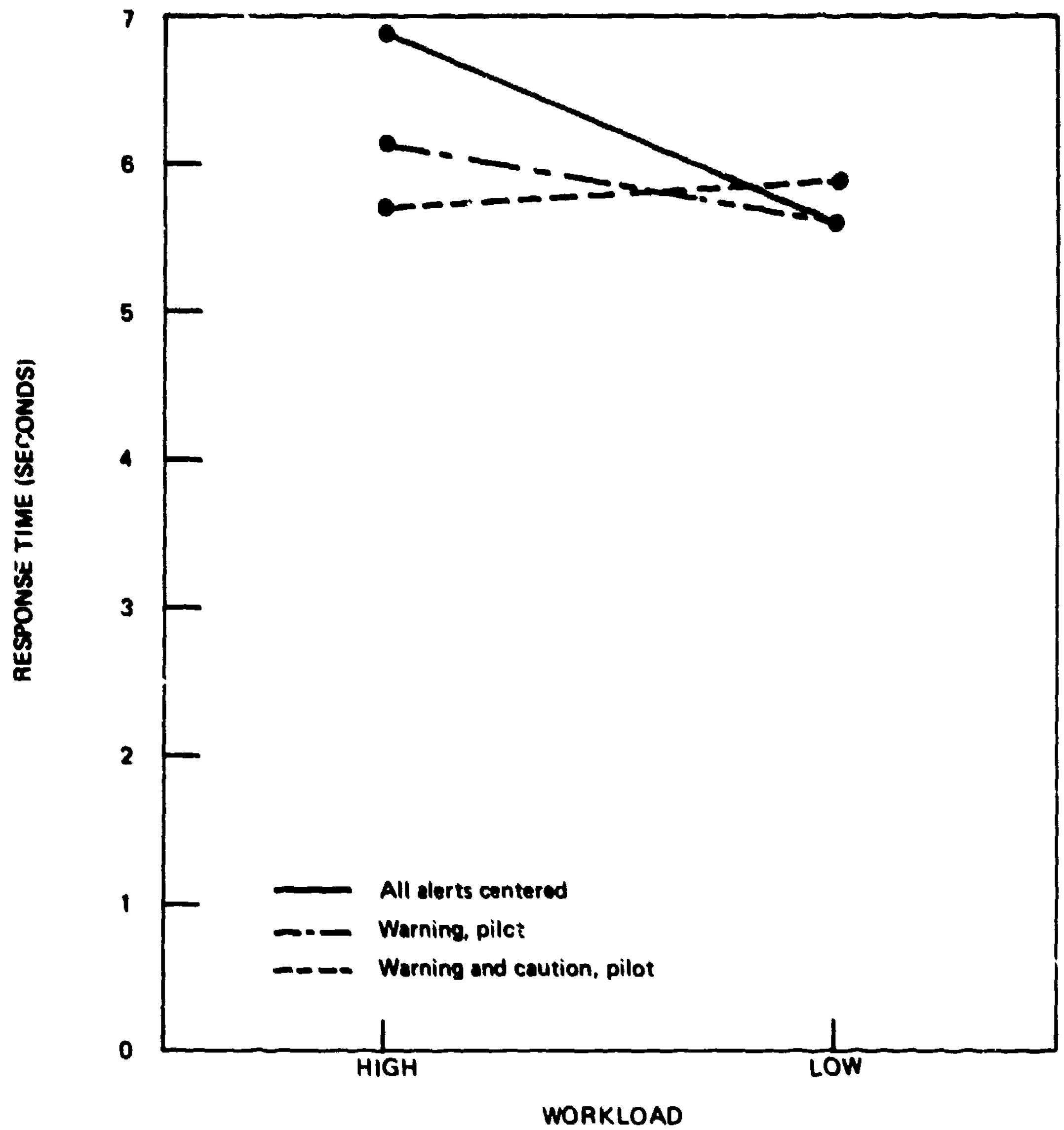


Figure 5.8.2.2-1. Mean Response Time as Function of Display Location and Workload

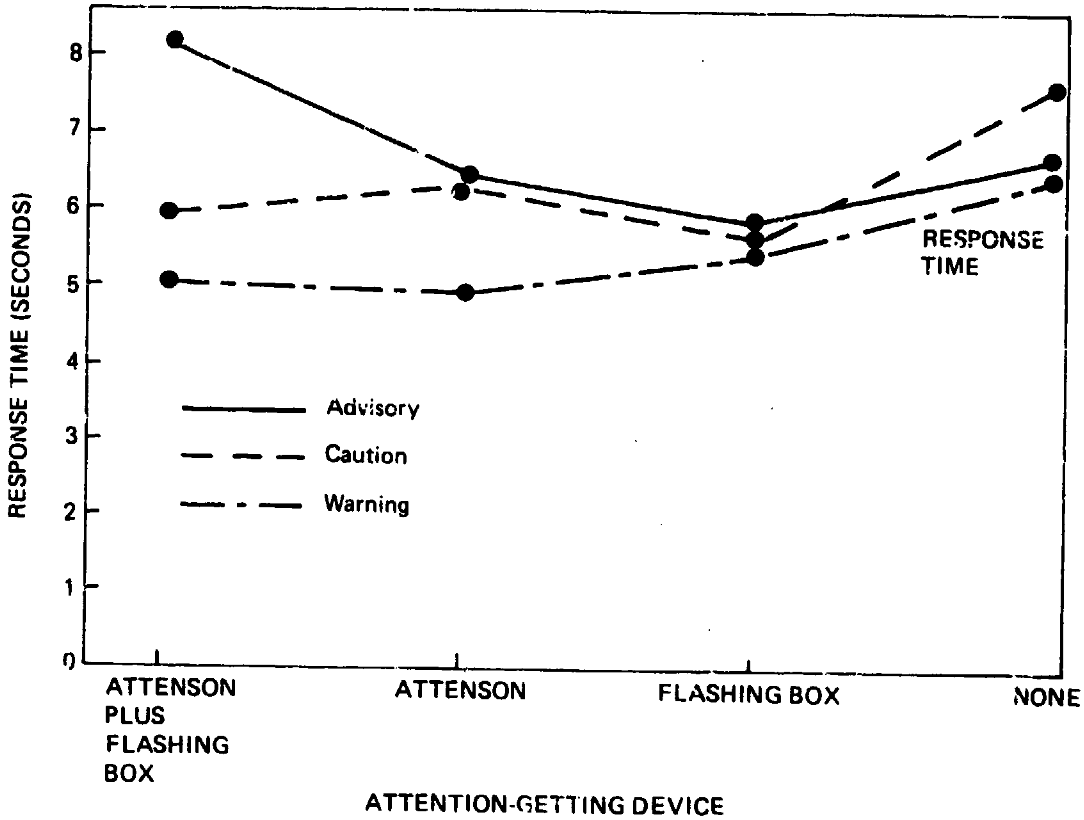


Figure 5.8.2.2-2. Test 2: Response Time as Function of Attention-Getting Device—All Alerts on Center Display

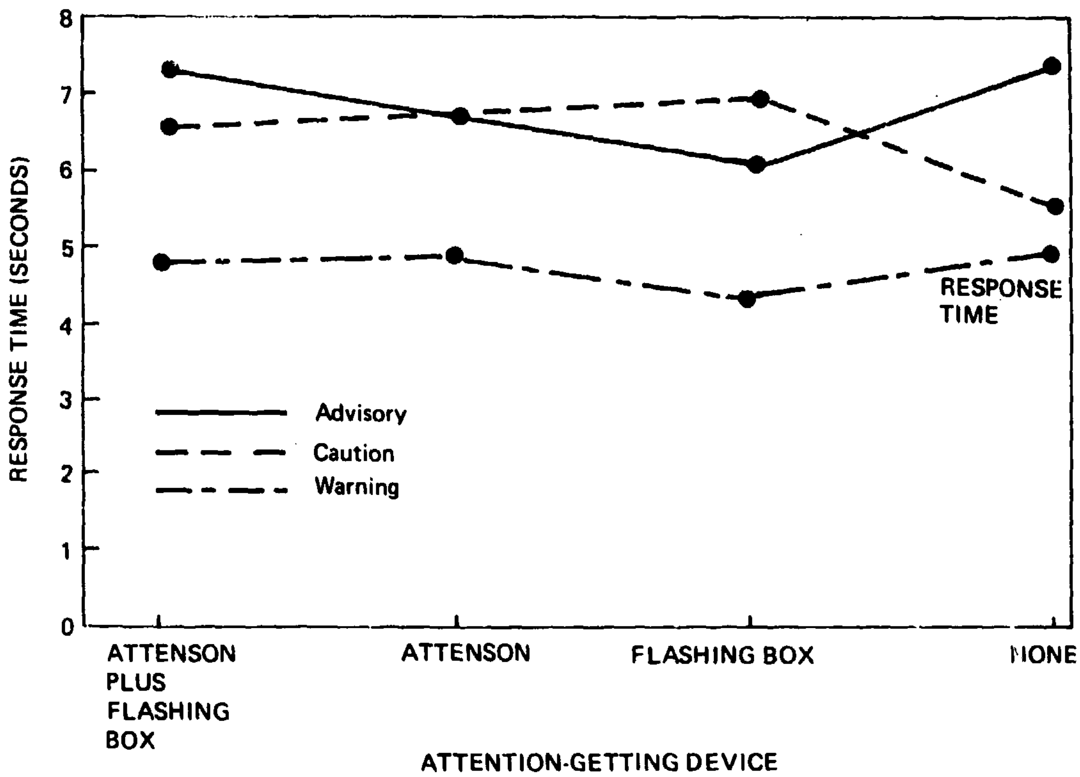


Figure 5.8.2.2-3. Test 2: Response Time as Function of Attention-Getting Device--Warnings on Pilot Display

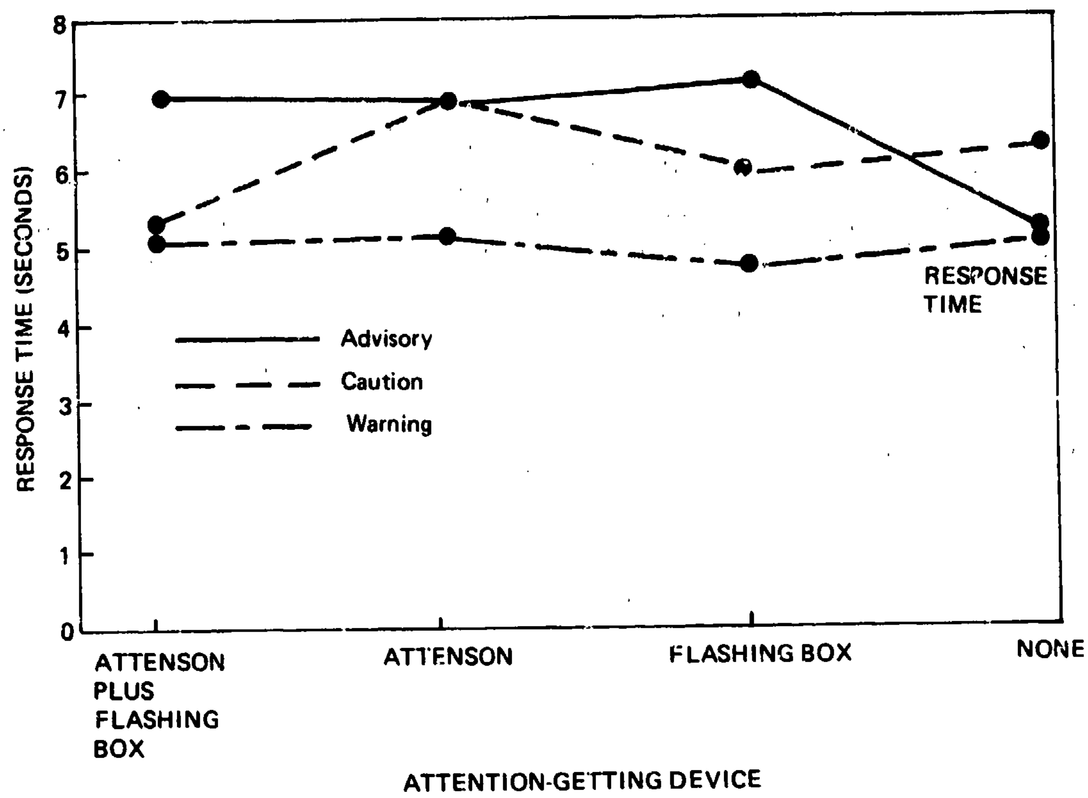


Figure 5.8.2.2-4. Test 2: Response Time as Function of Attention-Getting Device—Warning and Cautions on Pilot Display

accompanied by a flashing box and with cautions and warnings on different displays, the caution response times (6.3 sec, 6.38 sec, 6.66 sec, and 6.83 sec) were significantly slower than the warnings (4.87 sec, 4.67 sec, 4.68 sec, and 5.06 sec) even though the detection times were comparable.

When the attenson and flashing box were used together and all the alerts were presented on the same display, the mean response times for the cautions (5.89 sec and 5.23 sec) were not significantly different than those of the warnings (5.07 sec and 4.97 sec). The location of the display did not make a difference.

5.8.2.3 MISSED ALERTS

In an experiment of this type where the pilot is primed to detect and respond alerts, it is very difficult to produce conditions which will result in a large number of missed alerts. For test two there were a total of 2016 alerts out of which only 74 were missed (3.7% of the total). As in test one the attention getting quality of the alert proved to be the most significant contributer to the number of missed alerts: advisories which had neither an attenson or a flashing box were missed 6.5 percent of the time (44 out of 672); cautions and warnings which were presented without an attention-getter 25 percent of the time were missed 2.5 and 1.9 percent of the time respectively (17 and 13 out of 672). This difference which can be seen in Figure 5.8.2.3-1 was highly significant ($X^2 = 23.0$ $P < .001$). Table 5.8.2.3-1 and Figure 5.8.2.3-2 show that 88 percent of all missed alerts occurred when there was no attention-getting device (65 out of 74); 87 percent of the missed cautions and warnings (26 out of 30) occurred when there was no master visual attenson. No difference in the effectiveness of the attention-getters could be detected because of the extremely small numbers of missed alerts (inside dashed box on Table 5.8.2.3-1); however, the flashing box alone did produce 55 percent of the missed warnings and cautions (5 out of 9). The difference between the treatments was significant ($X^2 = 10.11$ $P < .02$). Neither the location of the display or the amount of pilot workload had a significant effect on the number of missed alerts.

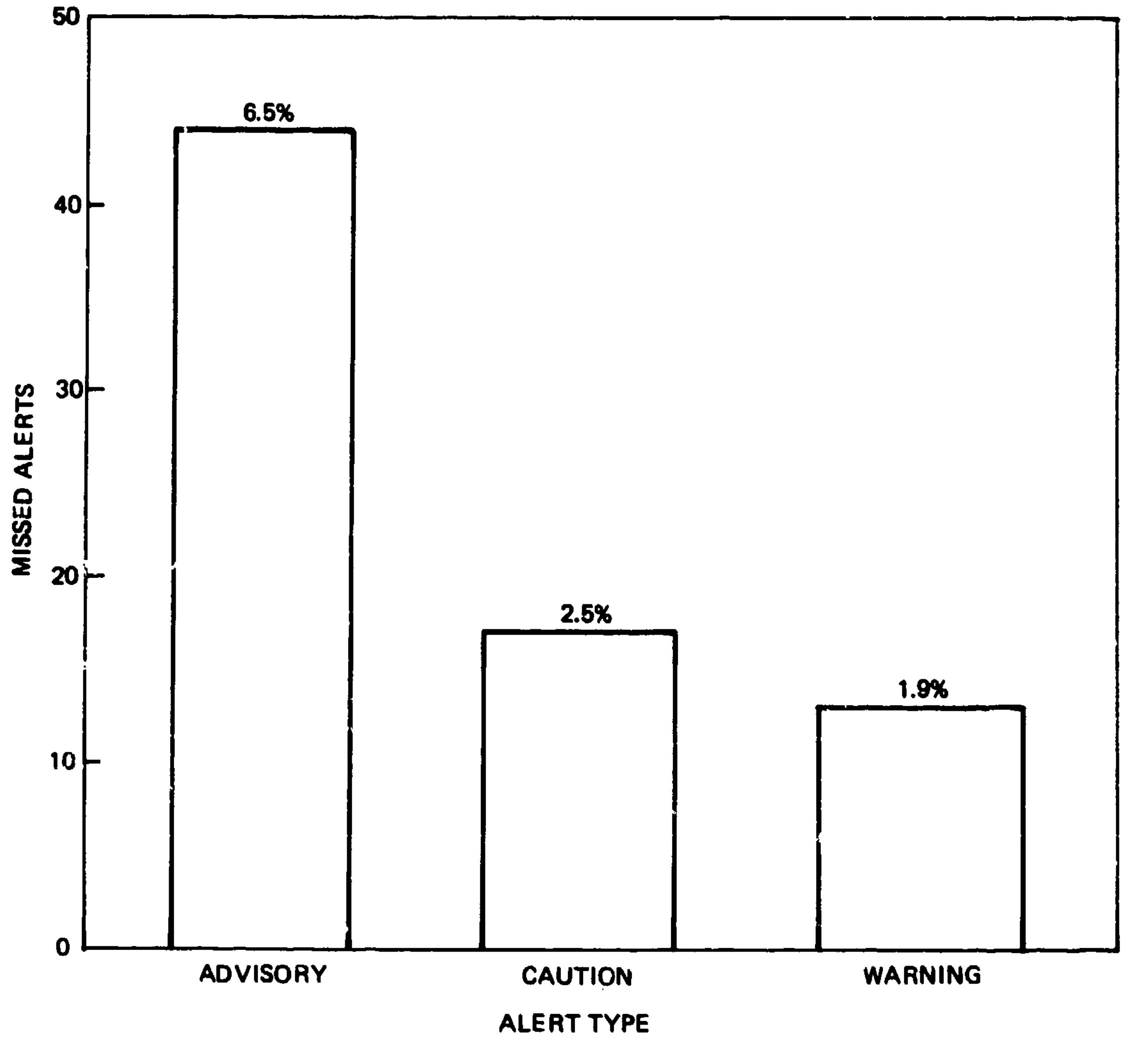


Figure 5.8.2.3-1. Test 2: Missed Alerts as Function of Alert Type

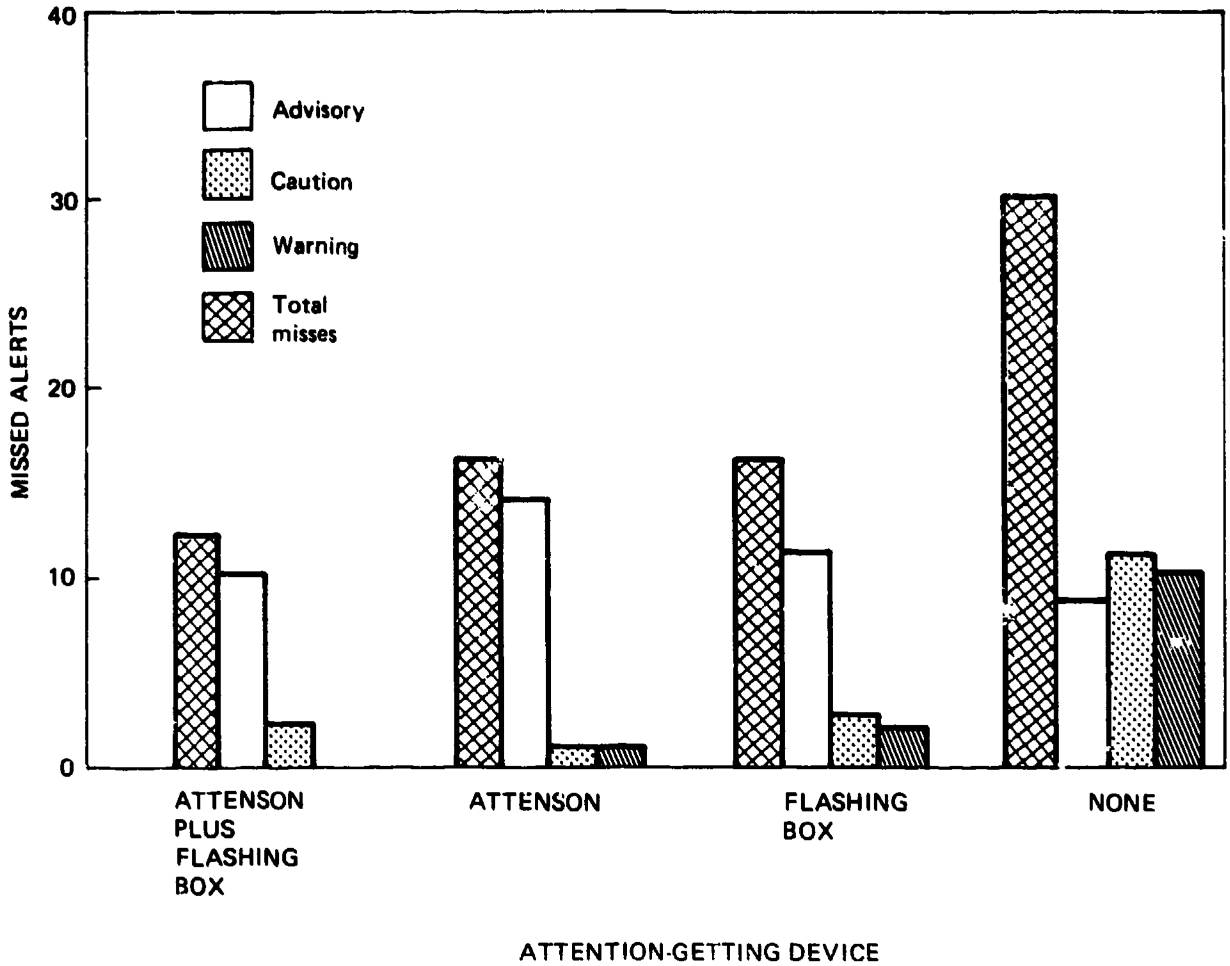


Figure 5.8.2.3-2. Test 2: Missed Alerts as Function of Attention-Getting Device

Table 5.8.2.3-1. Missed Alerts as Function of Attention-Getting and Alert Type

	Attenson plus flashing box	Attenson	Flashing box	None	Total
Advisory	10	14	11	9	44
Caution	2	1	3	11	17
Warning	0	1	2	10	13
Total	12	16	16	30	74

5.8.2.4 PILOT PREFERENCES

The questionnaire used for test two is included in Appendix B. The pilots preferred the master visual attenson (mean score 8.3) to either of the flashing box choices (5.8 for both) although they felt that the flashing box provided more information (8.9) than the master by itself (6.5). Four pilots commented that the master attenson should be combined with the box to provide the most information. Appendix E contains a compelation of the pilot comments.

The pilots showed a clear preference for presenting all the alerts in the center display (8.0) rather than warnings only (4.8) or cautions and warnings only (4.4).

When reporting the five features which they liked best, five of the seven pilots mentioned the master alert. The flashing box was mentioned four times; the concentration of all the alerts in one place and the three color system were mentioned three times. Three pilots disliked splitting up the alerts; this was the only disliked feature mentioned more than once. Four pilots recommended the addition of a cancel/recall function; two suggested that the flashing box should go steady after a certain time.

5.8.3 RESULTS OF DATA ANALYSIS FROM TEST 3

5.8.3.1 RESPONSE TIMES

The significance of differences in time required to respond to failure messages was evaluated by means of the Analysis of Variance Technique (ANOVA). The ANOVA summary table for Test 3 is shown in Table 5.8.3.1-1. The main effect attributed to level of auditory workload (concurrent vs. non-concurrent ATC communications) was significant at the .05 level. A significant interaction between auditory task loading and voice alert message format was also obtained ($p < .10$). Response time trends, based on average values across all subjects are plotted for each of these test parameters in Figures 5.8.3.1-1, -2, -3. No significant differences in mean response times were observed as a function of alerting mode (tone-no tone), voice alert message format (word-sentence), or ATC voice quality (male, female).

5.8.3.2 ATC RECOGNITION ACCURACY

Summary statistics for ATC recognition accuracy are presented in Table 5.8.3.2-1 and plotted in Figure 5.8.3.2-1. The significance of differences in total error frequencies across test conditions was evaluated using the Chi-Square (X^2) test for independence. Differences between low and high auditory workload levels were highly significant ($p < .01$). This effect was consistent across all combinations of alerting mode and voice message format. Although some evidence of differential masking of communications was observed as a function of controller voice quality (see Figures 5.8.3.1-1 and 5.8.3.1-2) this difference was not statistically significant.

5.8.3.3 ERROR ANALYSIS

In attempting to identify and respond verbally to ATC communications in a degraded auditory environment, pilots made several different types of errors. Recognition errors were classified and ranked according to degree of severity. These categories were defined as follows:

Readback Error: A "readback error" was scored when the test subject made an incorrect verbal response to an ATC advisory message. In an

Table 5.8.3.1-1. Test 3: Analysis of Variance Summary Table (Reaction Time Data)

Source	Sum of squares	Degrees of freedom	Mean square	F ratio
Main	3,610.90574	1	3,610.90574	53.24
Error	474.75826	7	67.82261	
Format	0.06891	1	0.06891	0.03
Error	18.81317	7	2.68760	
Tone	10.04080	1	10.04080	3.07
Error	22.88235	7	3.26891	
Format x tone	0.92991	1	0.92991	0.70
Error	9.26316	7	1.32331	
ATC timing	46.23613	1	46.23613	7.63*
Error	42.42328	7	6.06047	
Format x ATC timing	3.97268	1	3.97268	3.79*
Error	7.34018	7	1.04860	
Tone x ATC timing	1.86486	1	1.86486	1.31
Error	9.99882	7	1.42840	
Format x ATC timing	2.73488	1	2.73488	1.09
Error	17.48996	7	2.49857	
ATC voice	1.06763	1	1.06763	0.32
Error	23.44356	7	3.34908	
Format x ATC voice	0.12814	1	0.12814	0.18
Error	4.93393	7	0.70485	
Tone x ATC voice	0.86626	1	0.86626	0.44
Error	13.92780	7	1.09896	
Format x tone x ATC voice	2.04778	1	2.04778	2.76
Error	5.19007	7	0.74144	
ATC timing x ATC voice	0.00028	1	0.00028	0.00
Error	16.21882	7	2.31697	
Format x ATC timing x ATC voice	0.05241	1	0.05241	0.04
Error	9.12755	7	1.30394	
Tone x ATC timing x ATC voice	0.41976	1	0.41976	0.67
Error	4.41193	7	0.63028	
Format x tone x ATC timing x ATC voice	0.70954	1	0.70954	0.38
Error	12.94569	7	1.84938	

*Significant at the 0.10 level or better.

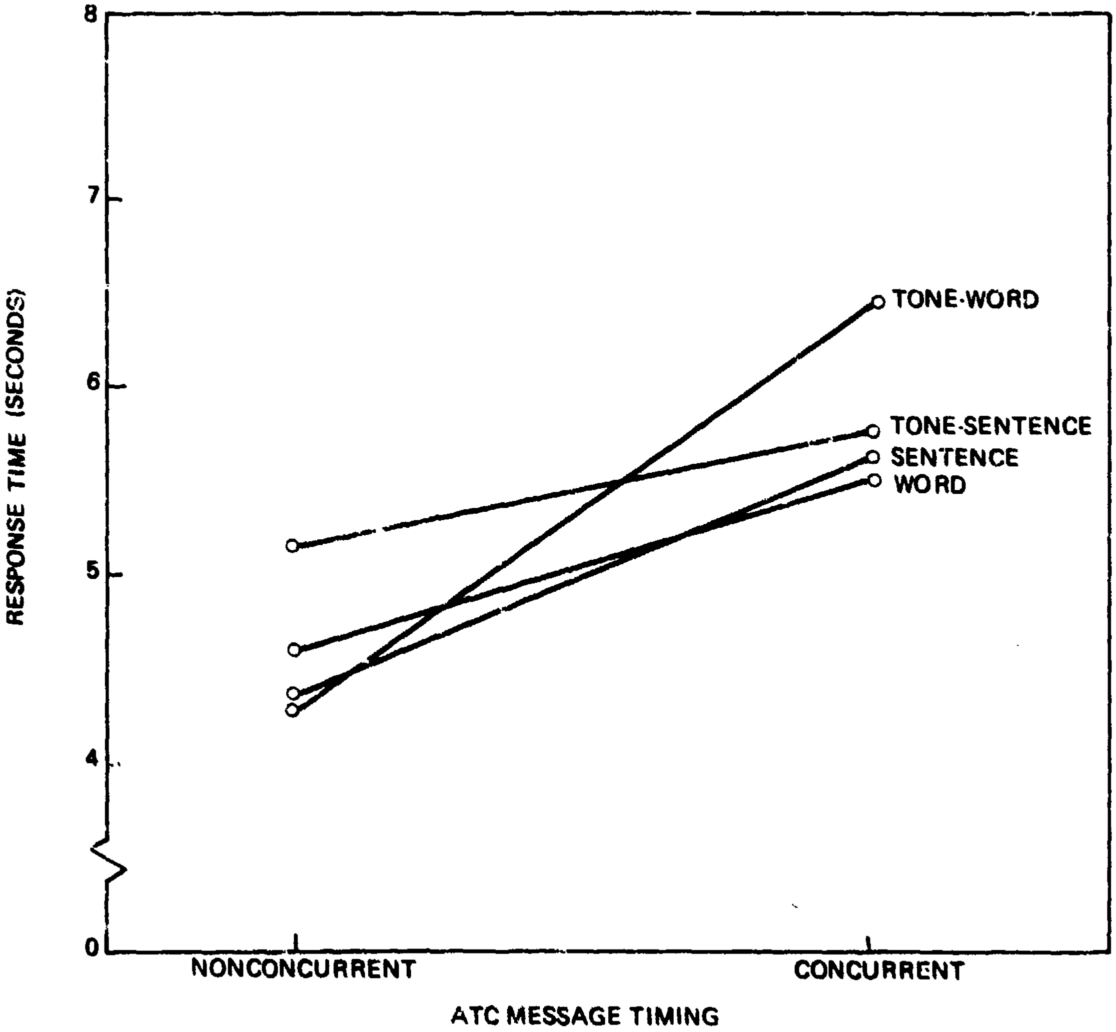


Figure 5.8.3.1-1. Response Time as a Function of Alert Message Format and ATC Message Timing: Male Controller Voice

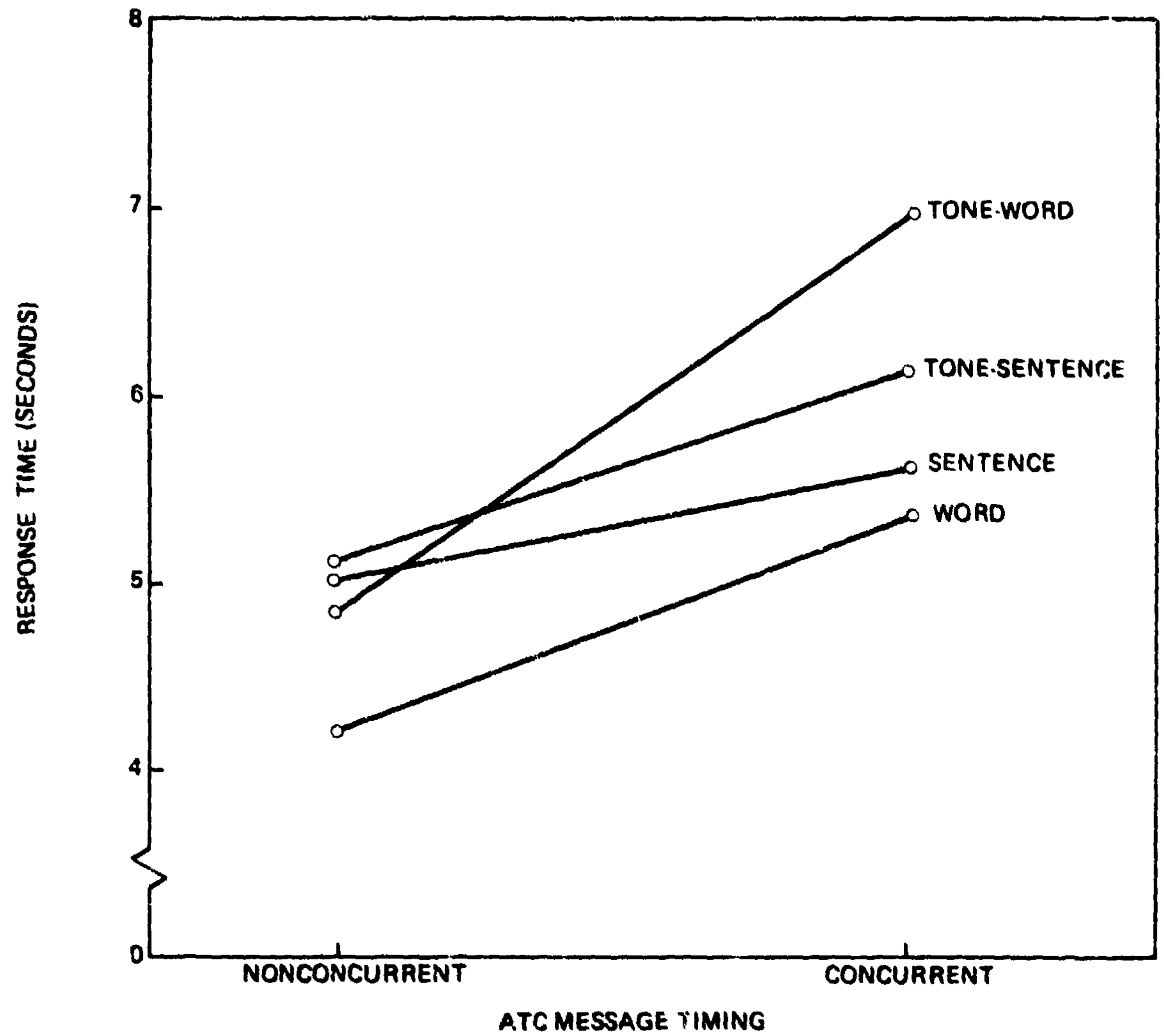


Figure 5.8.3.1-2. Response Time as a Function of Alert Message Format and ATC Message Timing: Female Controller Voice

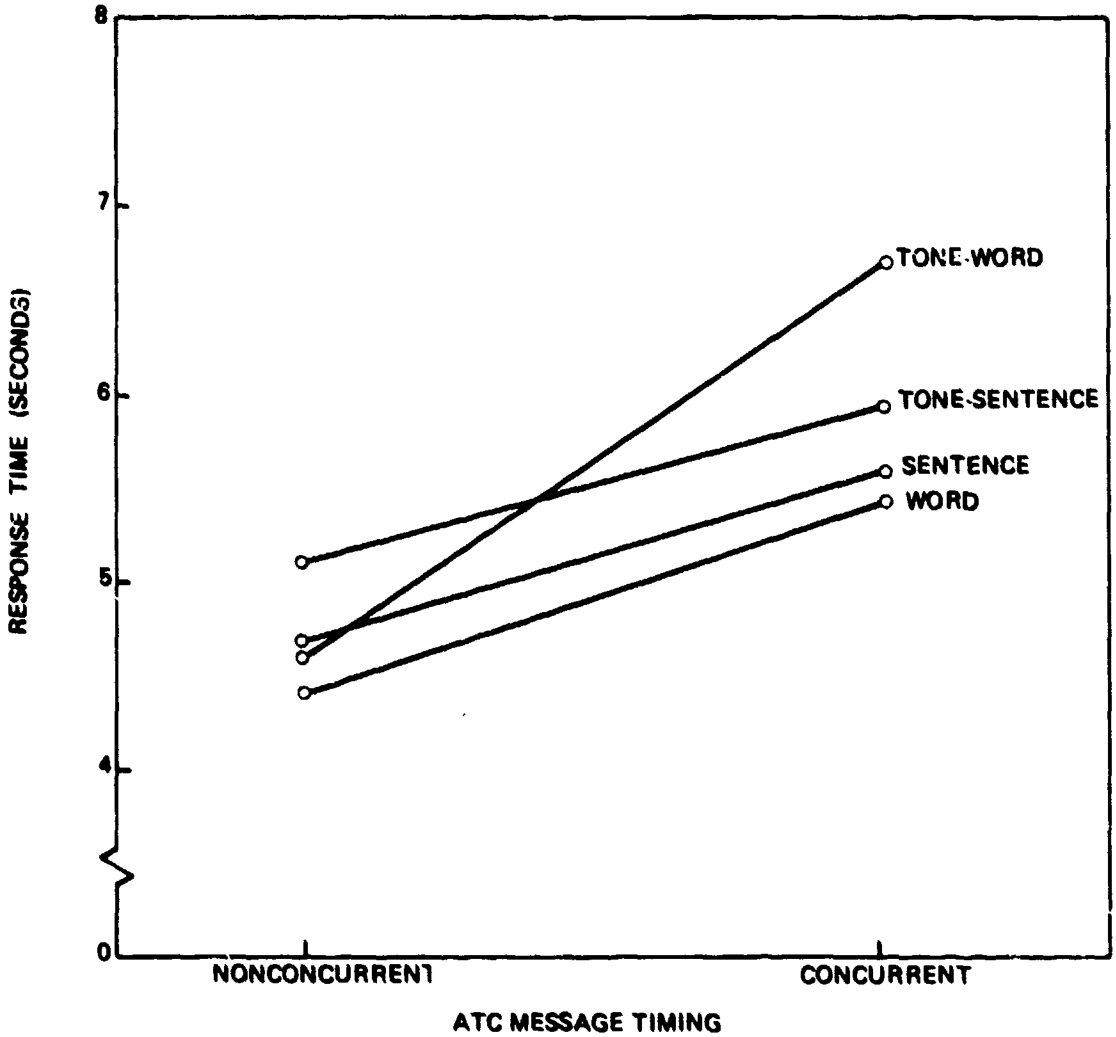


Figure 5.8.3.1-3. Response Time as a Function of Alert Message Format and ATC Message Timing: Combined Data for Male and Female Controller Voices

Table 5.8.3.2-1. Summary Statistics for Percent ATC Recognition Accuracy as a Function of Alert Message Format and ATC Message Timing: Male and Female ATC Voices

Message format	ATC message timing	
	Nonconcurrent	Concurrent
Word	6.2	87.5
Sentence	0.0	81.2
Tone-word	12.5	75.0
Tone-sentence	6.2	94.0

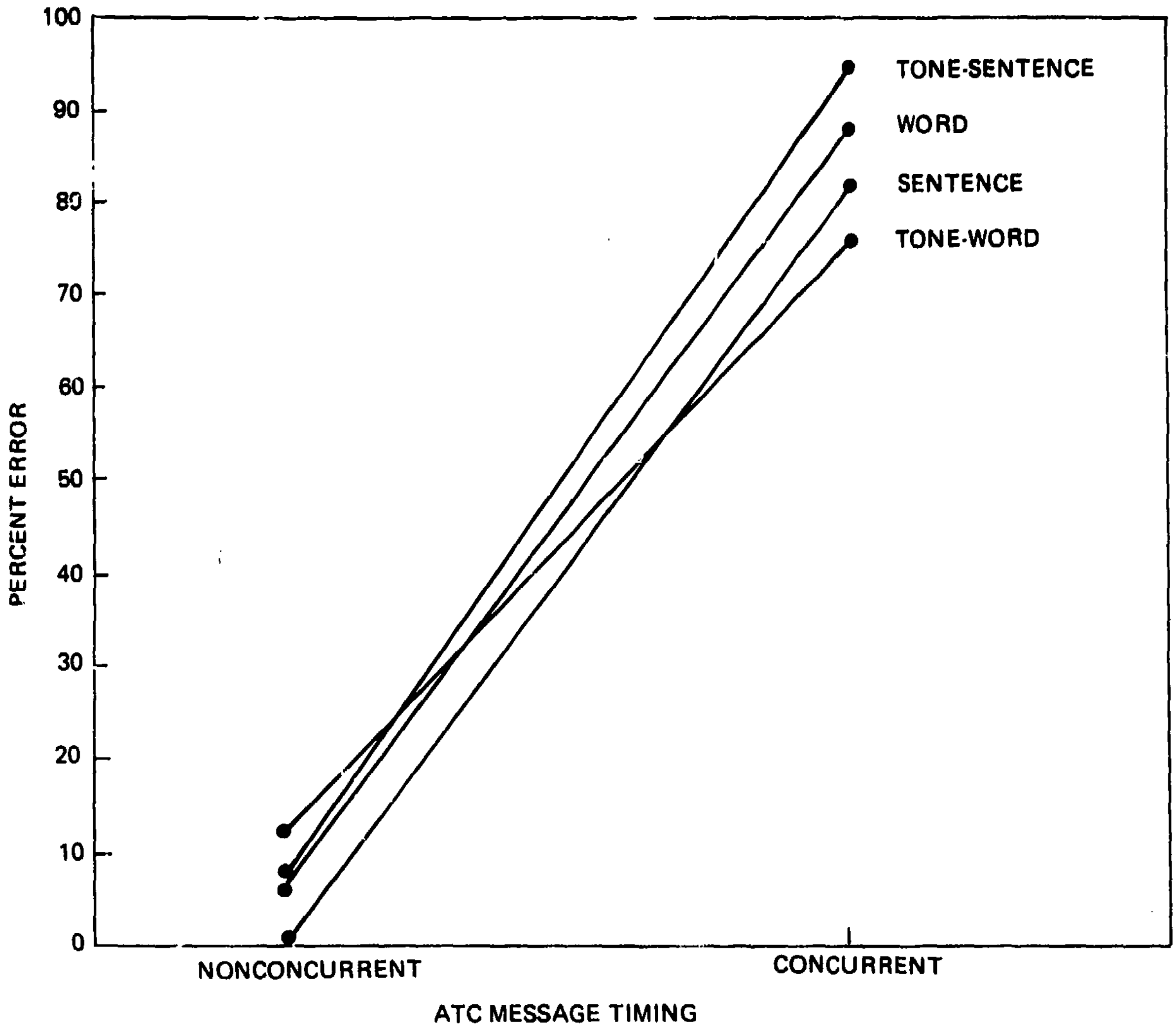


Figure 5.8.3.2-1. Percent Error on the ATC Recognition Task as a Function of Alert Message Format and ATC Message Timing: Combined Data for Male and Female Controller Voices

operational environment, this is the most serious type of error, since misinterpretation of an ATC instruction might lead to an incorrect action by the flight crew.

Omission Error: An "omission error" was scored when a test subject failed to acknowledge an ATC transmission. This type of error is less critical since the controller would normally repeat the message if no acknowledgement is received.

Repeat Error: A "repeat error" was scored when the test subject requested a repeat of the ATC communication. This is considered the least serious type of error.

Total error frequencies for each of the experimental conditions were evaluated according to this classification system.

Chi-square tests for independence were performed for each of the test parameters to determine if the distribution of error types varied as a function of alerting mode, message format or ATC voice quality. These data are summarized in Figures 5.8.3.3-1 through 5.8.3.3-4 and contingency Tables 5.8.3.3-1 through 5.8.3.3-4.

5.8.3.4 TRACKING TASK ACCURACY

Post-alert control deviations provided no evidence of differential disruptive effects as a function of the auditory system characteristics under evaluation. As anticipated, accuracy of aircraft control appears to be a relatively insensitive measure of auditory task loading. Based on test observations, it must be assumed that distracting effects on visual tasks are negligible when alert information is presented exclusively by means of auditory stimuli.

5.8.3.5 PILOT PREFERENCES

The distribution of pilot preferences for alerting mode and message format alternatives are summarized in Tables 5.8.3.5-1, -2. A large majority (87%) of pilots participating in Test III felt that the presence of a precursor tone

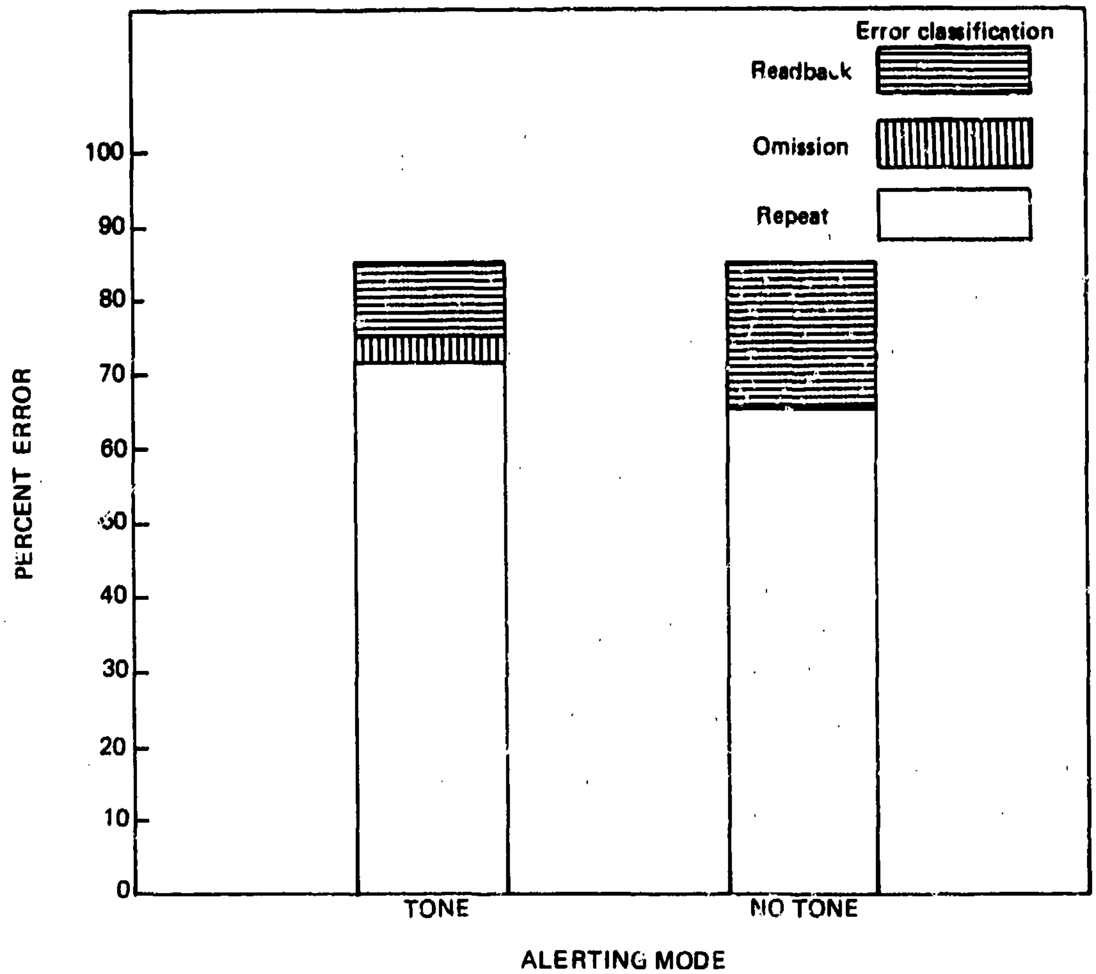


Figure 5.8.3.3-1. Distribution of ATC Recognition Error Types With and Without Precursor Tone: Concurrent ATC Communications

Table 5.8.3.3-1. Error Analysis Contingency Table for ATC Recognition Task: Concurrent ATC Communications

	Error classification			
Alerting mode	Repeat	Omission	Readback	Total
Tone	23	1	3	27
No tone	21	6	0	27
Total	44	7	3	54

X^2 test for independence = 3.64*

X^2 alerting mode marginal total = 0.00

X^2 error classification marginal totals = 56.78, p = 0.001

*Note: Computed using Yates correction for continuity.

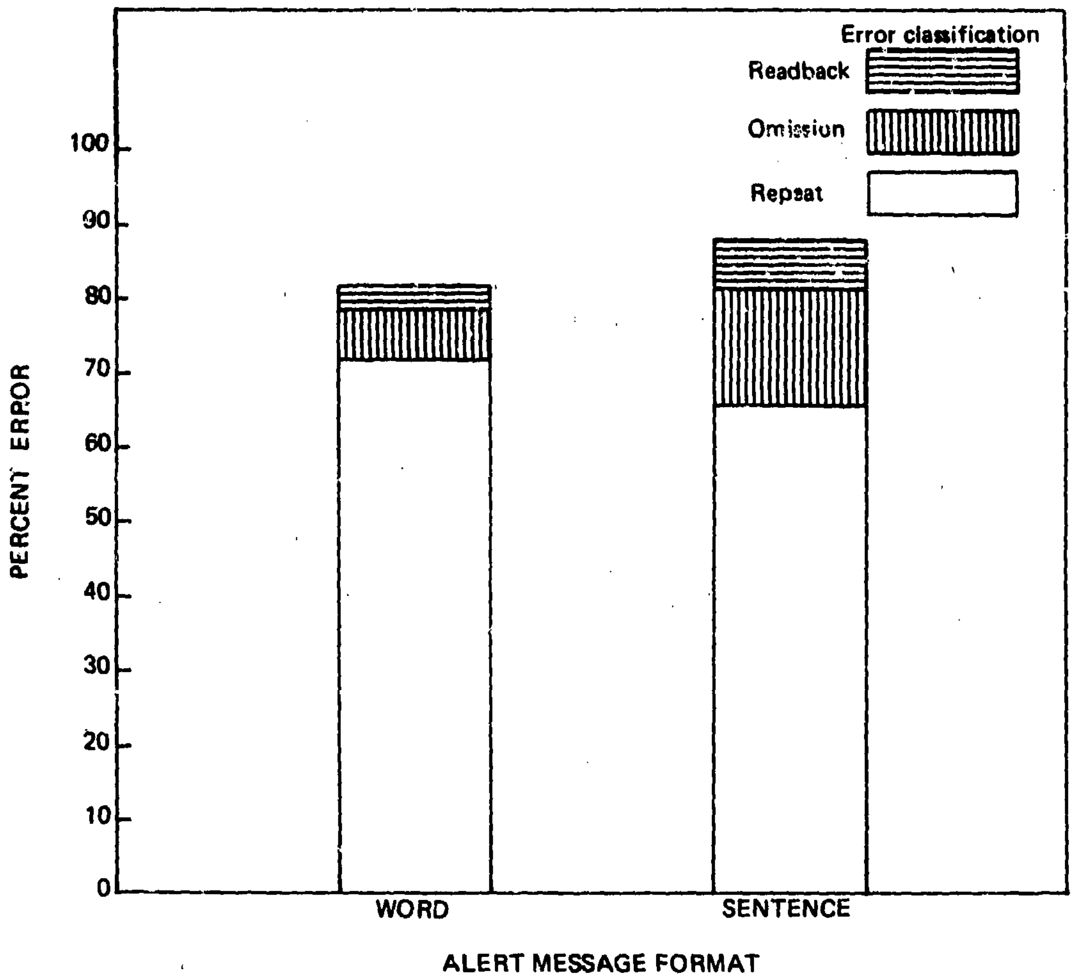


Figure 5.8.3.3-2. Distribution of ATC Recognition Error Types Across Alert Message Formats: Concurrent ATC Communications

Table 5.8.3.3-2. Error Analysis Contingency Table for ATC Recognition Task: Concurrent ATC Communications

Error classification				
Message format	Repeat	Omission	Readback	Total
Word	23	2	1	26
Sentence	21	5	2	28
Total	44	7	3	54

X^2 test for independence = 0.59*
X^2 message format marginal totals = 0.08
X^2 error classification marginal totals = 56.78, p = 0.001

*Note: Computed using Yates correction for continuity.

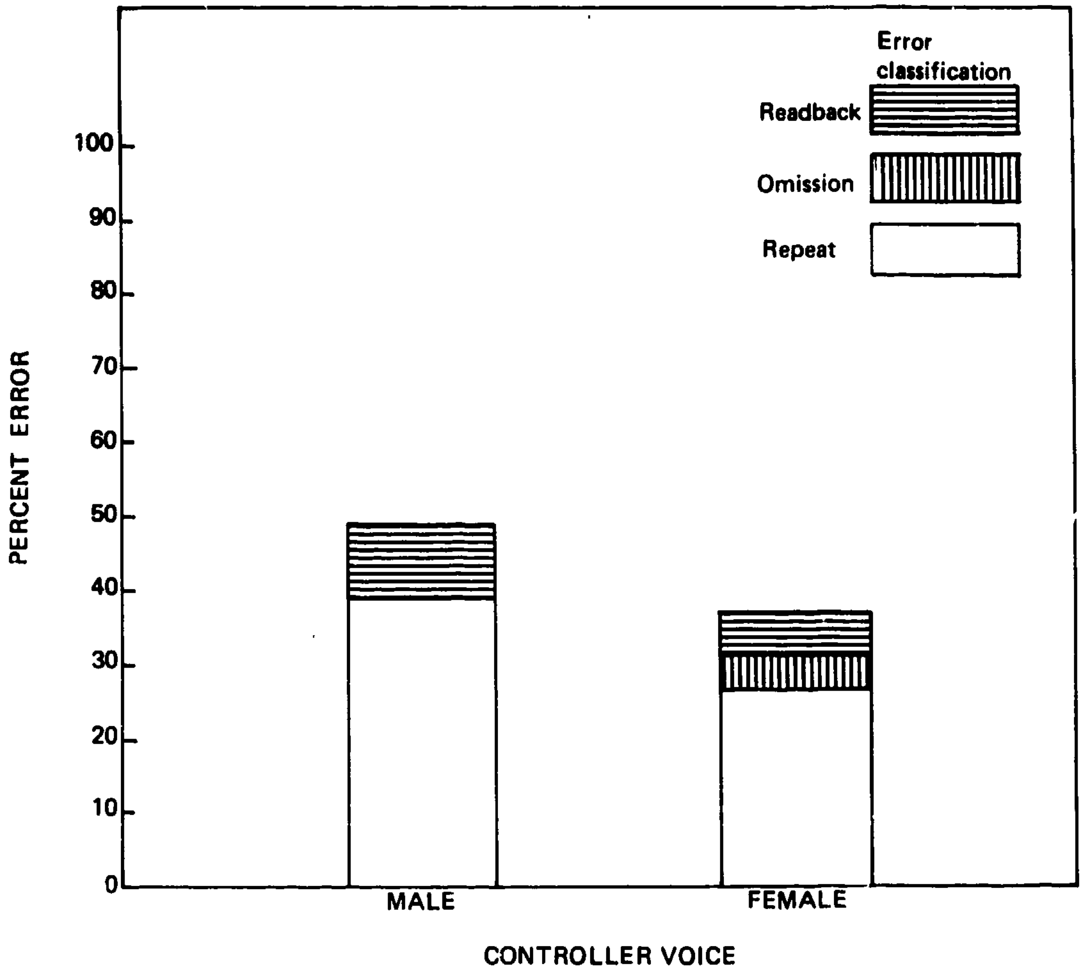


Figure 5.8.3.3-3. Distribution of ATC Recognition Error Types Across Male and Female Controller Voices: Concurrent ATC Communications

Table 5.8.3.3-3. Error Analysis Contingency Table for ATC Recognition Task: Concurrent ATC Communications

Controller voice	Error classification: Repeat	Omission	Readback	Total
Male	25	6	0	31
	78%	19%	0%	97%
Female	19	1	3	23
	59%	3%	9%	72%
Total	44	7	3	54

X^2 test for independence = 3.30*

X^2 controller voice marginal totals = 1.19

X^2 error classification marginal totals = 56.78, p = 0.001

*Note: Computed using Yates correction for continuity.

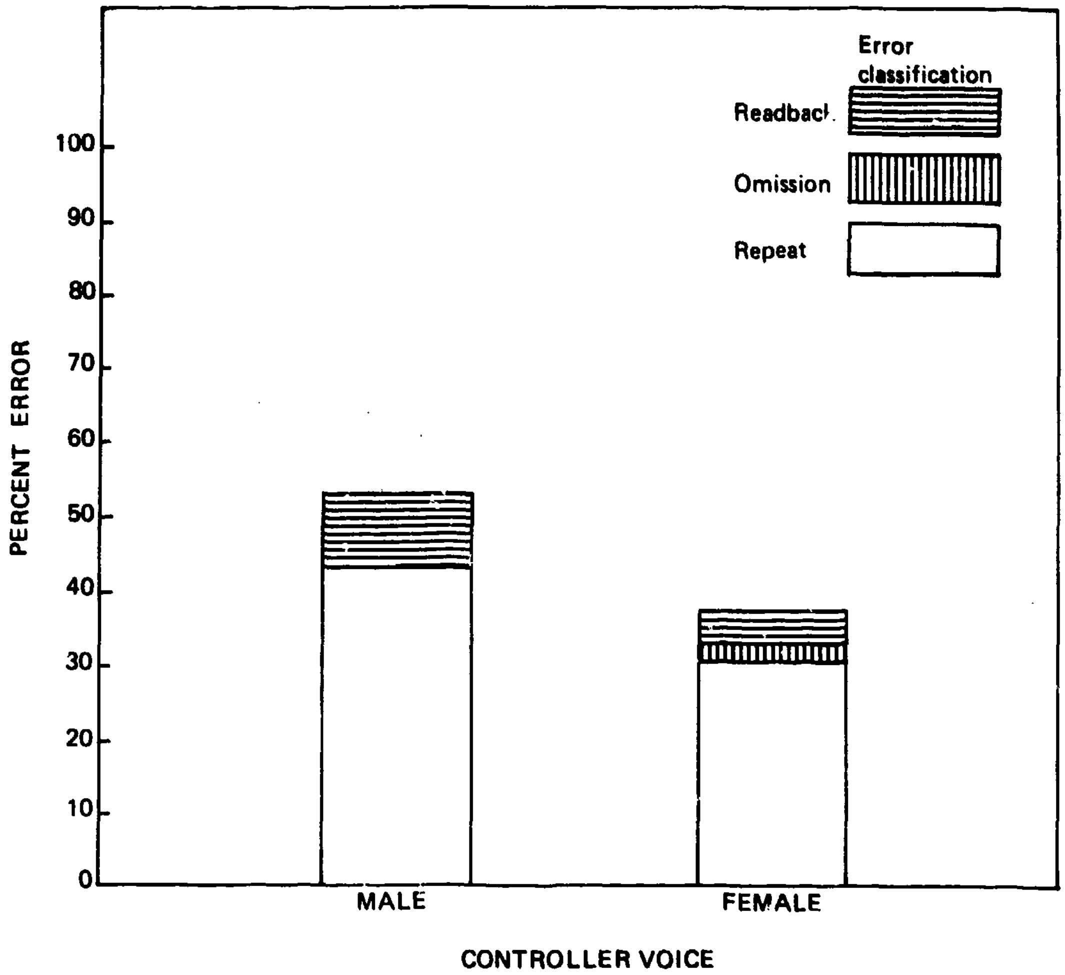


Figure 5.8.3.3-4. Distribution of ATC Recognition Error Types Across Male and Female Controller Voices; Concurrent and Nonconcurrent ATC Communications

Table 5.8.3.3-4. Error Analysis Contingency Table for ATC Recognition Task: Frequencies Combined for Concurrent and Nonconcurrent ATC Communications

Controller voice	Error classification: Repeat	Omission	Readback	Total
Male	28	6	0	34
Female	20	1	3	24
Total	48	7	3	58

X^2 test for independence = 3.34*

X^2 controller voice marginal totals = 1.72

X^2 error classification marginal totals = 64.18, p = 0.001

*Note: Computed using Yates correction for continuity.

Table 5.8.3.5-1. Pilot Preferences for Tone-Voice Options (n = 8)

	Voice only		About equal		Tone-voice	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Which alerting mode was most effective in getting your attention?	0	0.00	1	12.5	7	87.5
Which alerting mode would you prefer for a cockpit warning system?	1	12.5	0	0.0	7	87.5

Table 5.8.3.5-2. Pilot Preferences for Voice Message Formats (n = 8)

	Word phrase		Complete sentence		About equal	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Which voice alert format was easiest to identify when presented simultaneously with an ATC message?	6	75.00	2	25.00	0	0.00
Which voice alert format caused the least difficulty in identifying concurrent ATC messages?	5	62.00	2	25.00	1	12.50
Which voice alert format would you prefer for a cockpit alerting system?	5	62.50	3	37.50	0	0.00

enhanced the attention-getting value of the voice alert message. An equal number of pilots selected the tone-voice mode as the preferred option for a cockpit alerting system. When asked to indicate their preferences for voice message format, most pilots (75%) felt that the short word or phrase structure would be easiest to identify and interpret. A somewhat smaller majority (62%) felt that the word/phrase format would tend to minimize masking effects on conflicting ATC communications.

Judgements of overall effectiveness favored the word/phrase format over the complete sentence format by a margin of 62.5% to 37.5%. It should be noted that some pilots selected the sentence format because it provided a more complete description of the failure condition rather than improved intelligibility resulting from language context.

5.8.4 RESULTS OF DATA ANALYSIS FROM TEST 4

5.8.4.1 RESPONSE TIMES

Results of the analysis of variance for Test 4 response time data are summarized in Table 5.8.4.1-1. As in Test 3, the presence of concurrent ATC communications had a pronounced effect on the time required to identify and respond to alert messages. The F-ratio attributed to the auditory workload main effect was highly significant ($p < .01$). The data plotted in Figure 5.8.4.1-1 are response times averaged across test subjects for the low and high auditory workload conditions. The plots represent combined data for both levels of tracking task difficulty. Increases in mean response times associated with the high auditory workload condition ranged from 1.37 to 2.21 seconds for alerting modes involving voice annunciations. Inspection of Figure 5.8.4.1-1 shows a relatively minor increase in response time of .69 seconds for the tone-visual alerting mode. It should be noted that this apparent interaction between alerting mode and auditory task loading was not statistically significant. This differential effect is reported here because the results tend to correlate highly with ATC recognition accuracy data.

The level of turbulence (tracking task difficulty) did not influence the time required to identify and respond to alert messages. No significant differences in response time were noted as a result of the first or second order interactions between alerting modes and levels of secondary task loading.

5.8.4.2 ATC RECOGNITION ACCURACY

The level of visual workload had no measurable effect on the pilot's ability to detect and interpret ATC communications. No evidence of significant interactions between visual task loading and other test parameters was observed in the ATC recognition accuracy data.

Summary statistics for the ATC recognition accuracy data are presented in Table 5.8.4.2-1 and overall error percentages are plotted in Figure 5.8.4.2-1. Since tracking task difficulty did not interact with other test variables, the values presented represent combined data for both levels of visual workload.

Table 5.8.4.1-1. Test 4: Analysis of Variance Summary Table (Reaction Time Data)

Source	Sum of squares	Degrees of freedom	Mean square	F ratio
Main	5865.25648	1	5865.25648	186.23
Error	220.46555	7	31.49508	
Mode	23.57039	3	7.85680	2.11
Error	78.06948	21	3.71759	
ATC timing	81.60032	1	81.60032	38.77*
Error	17.73404	7	2.10486	
Mode x ATC timing	12.45370	3	4.15123	0.99
Error	87.69814	21	4.17610	
Turbulence	0.28313	1	0.28313	0.12
Error	15.95935	7	2.27991	
Mode x turbulence	6.19765	3	2.06588	0.57
Error	75.77186	21	3.27991	
ATC timing x turbulence	5.31380	1	5.31380	2.30
Error	16.14625	7	2.30661	
Mode x ATC timing x turbulence	6.08241	3	2.02747	1.19
Error	35.72765	21	1.70132	

*Significant at the 0.10 level or better.

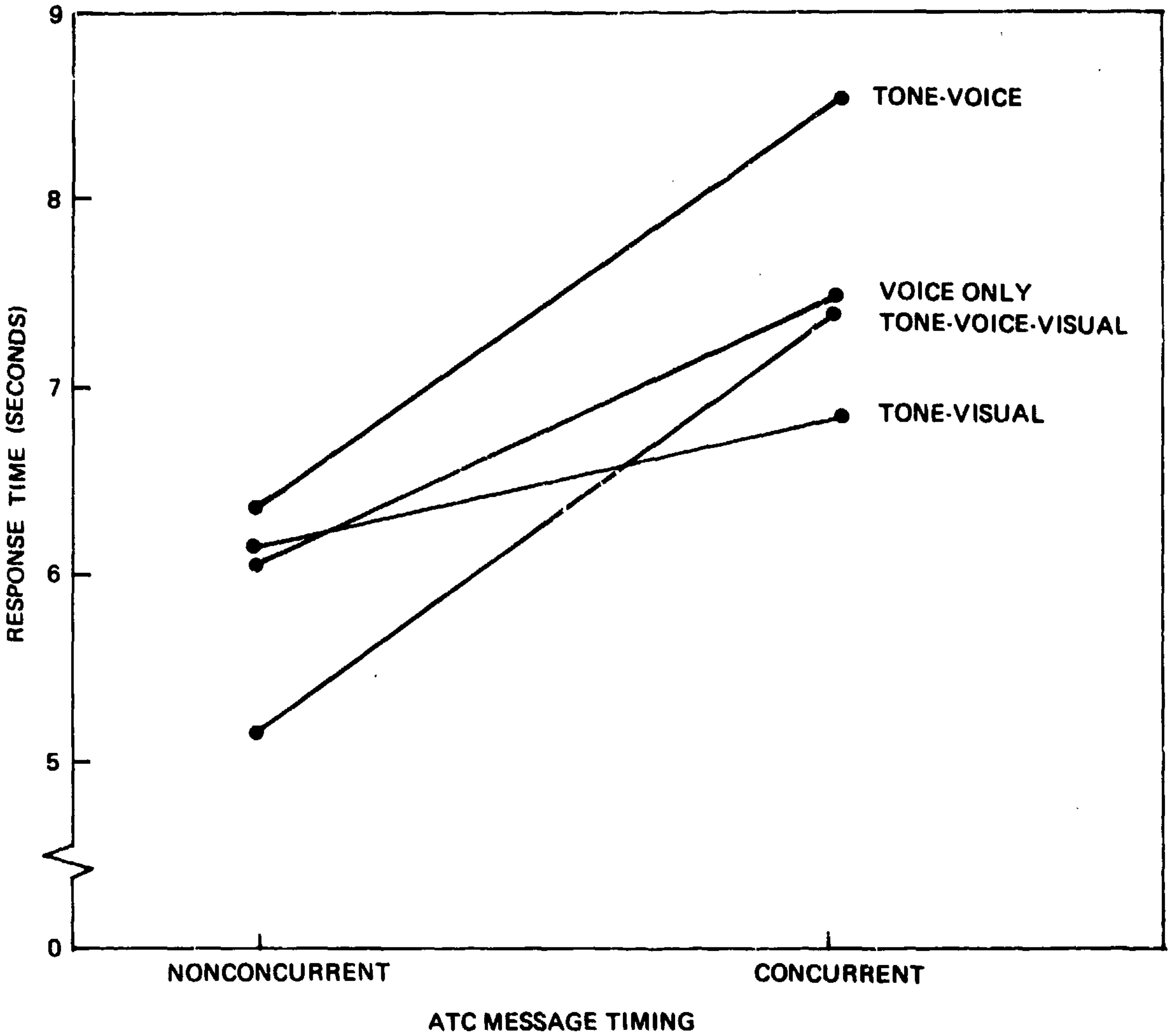


Figure 5.8.4.1-1. Response Time as a Function of Alerting Mode and ATC Message Timing; Combined Data for High- and Low-Turbulence Levels

Table 5.8.4.2-1. Summary Statistics for Percent ATC Recognition Accuracy as a Function of Alerting Mode and ATC Message Timing: High and Low Turbulence Level

	ATC message timing	
Alerting mode	Nonconcurrent	Concurrent
Tone-visual	19	31
Tone-voice	12	69
Voice only	19	75
Tone-voice-visual	0	56

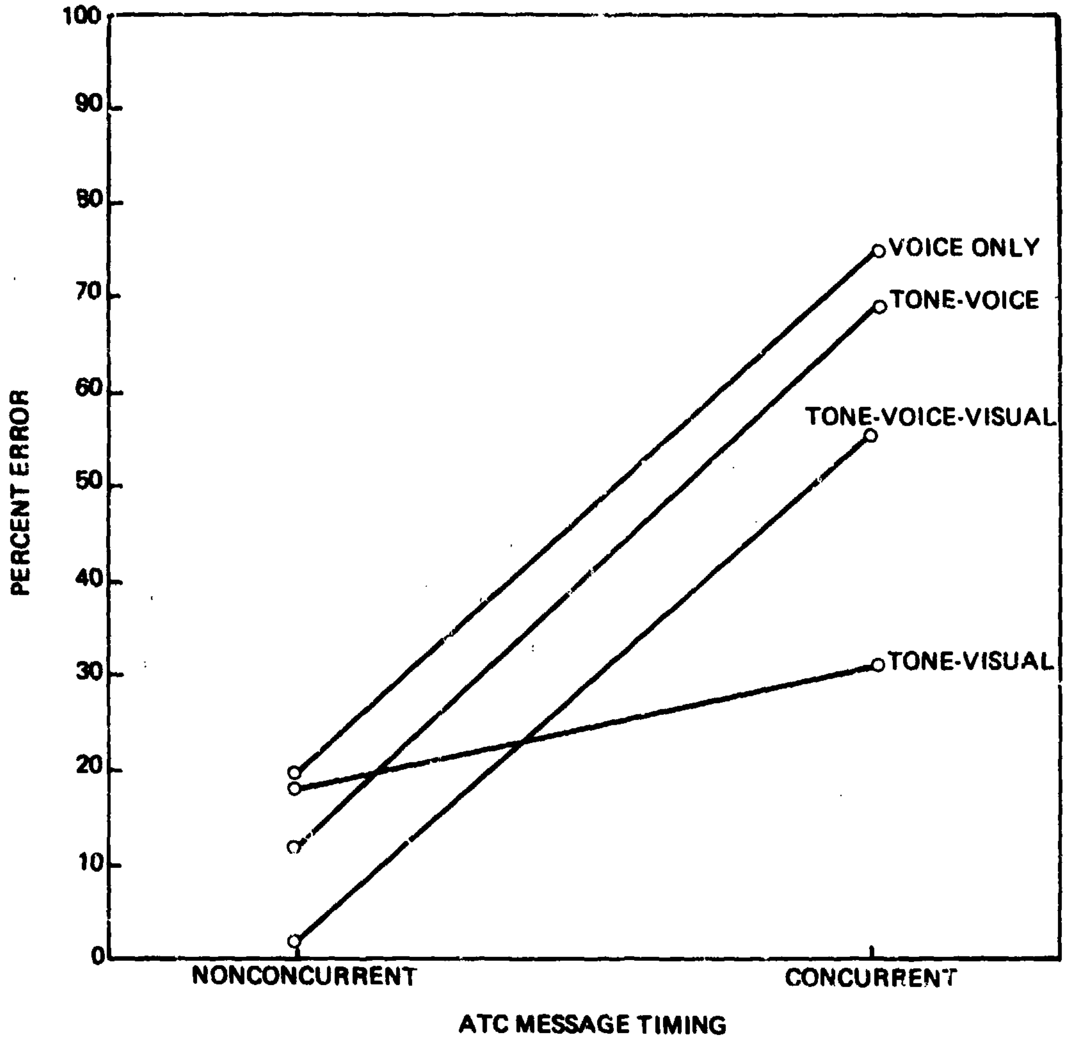


Figure 5.8.4.2-1. Percent Error on the ATC Recognition Task as a Function of Alerting Mode and ATC Message Timing: Combined Data for High and Low Turbulence Levels

The significance of differences in error frequencies was evaluated by means of the Chi-Square (X^2) statistic. As in Test 3, increases in error rates associated with concurrent ATC communications were found to be highly significant ($p < .01$). Inspection of Figure 5.8.4.2-1 reveals that this effect was essentially constant across all alerting modes including a voice alert message component. The increase in auditory workload resulted in a concomitant increase in the proportion of ATC recognition errors of approximately 56% for all voice alert modes. Introduction of simultaneous ATC communications resulted in a relatively small increase of 12% in the error rate for the tone-visual mode.

5.8.4.3 ERROR ANALYSIS

Total error frequencies for each of the experimental conditions were broken down according to the classification system described in Section 5.8.3.3, readback, omission and repeat errors.

Chi-Square tests for independence were performed to determine if the distribution of error types varied as a function of alerting mode. These data are summarized in Figure 5.8.4.3-1 and contingency Table 5.8.4.3-1.

The Chi-Square tests for significance of differences between error classification marginal totals yielded results that are generally consistent with the findings in test 3. Requests for message repeat ("say again" responses) represent the largest proportion of the total recognition errors.

5.8.4.4 TRACKING TASK ACCURACY

Localizer, roll and glideslope deviation performance measures recorded during the post-alert segment were analyzed by means of the analysis of variance. The ANOVA results for localizer deviation data are summarized in Table 5.8.4.4-1. The alerting mode main effect and mode by visual workload interaction were statistically significant ($p < .01$). Mean square localizer deviations are plotted in Figures 5.8.4.4-1, -2, -3 showing error trends as a function of alerting mode, turbulence level, and ATC message timing. These graphs show that alerts presented by means of the combined tone-voice-visual

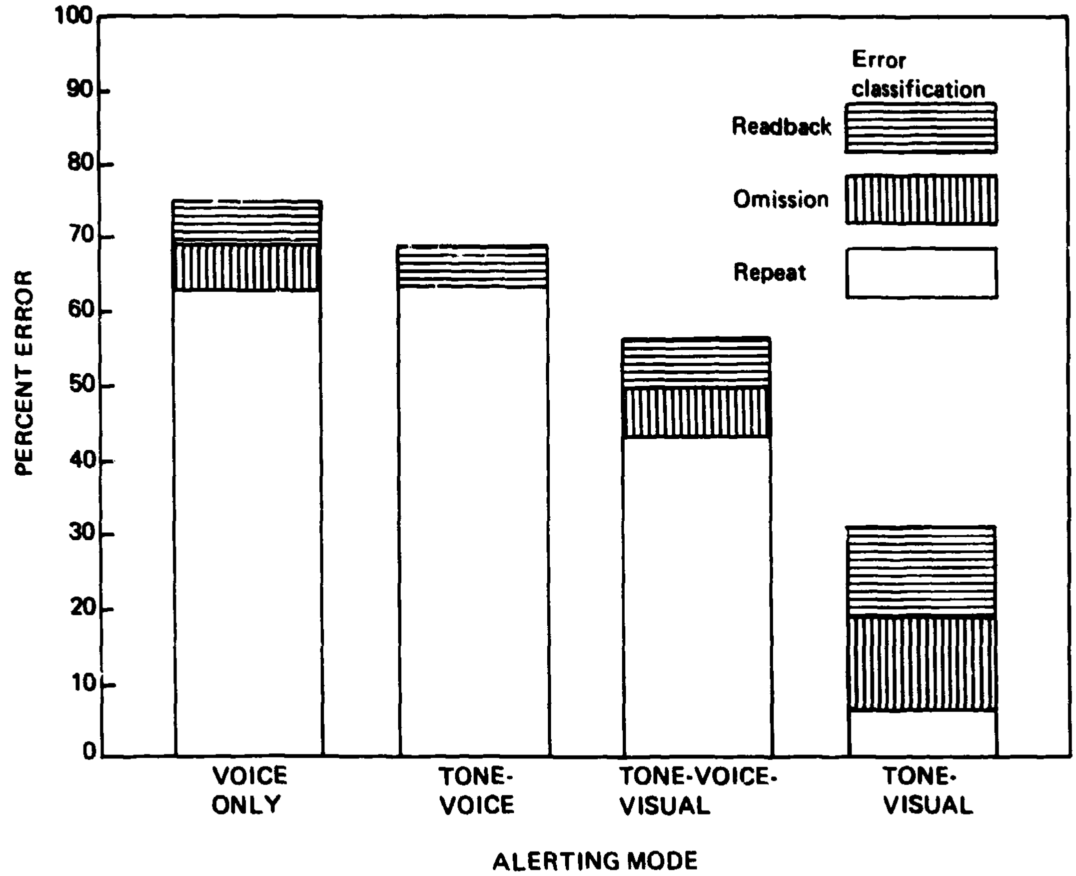


Figure 5.8.4.3-1. Distribution of ATC Recognition Error Types Across Alerting Modes: Concurrent ATC Communications

Table 5.8.4.3-1. Error Analysis Contingency Table for ATC Recognition Task: Concurrent ATC Communications

	Error classification			
Alerting mode	Repeat	Omission	Readback	Total
Tone-visual	1	2	2	5
Tone-voice	10	0	1	11
Voice only	10	1	1	12
Tone-voice-visual	7	1	1	9
Total	28	4	5	37

X^2 test for independence = 5.04*

X^2 alerting mode marginal totals = 3.11

X^2 error classification marginal total = 39.90, p = 0.001

*Computed using Yates correction for continuity

Table 5.8.4.4-1. Test 4: Analysis of Variance Summary Table (Localizer Deviation Data: Postalert Segment)

Source	Sum of squares	Degrees of freedom	Mean square	F ratio
Main	394002.53352	1	394002.53352	37.29
Error	73951.95251	7	10564.56464	
Mode	137267.11506	4	34316.77877	2.68*
Error	358599.64611	28	12807.13022	
ATC timing	4032.06267	1	4032.06267	0.16
Error	180143.91915	7	25734.84559	
Mode x ATC timing	48520.97820	4	12130.24455	0.99
Error	343036.44415	28	12251.30158	
Turbulence	305602.75200	1	305602.75200	36.18*
Error	59132.86758	7	8447.55251	
Mode x turbulence	129942.92826	4	32485.73207	2.62*
Error	347239.74780	28	12401.41956	
ATC timing x turbulence	3470.76820	1	3470.76820	0.13
Error	188979.35758	7	26997.05108	
Mode x ATC tim' x turbulence	52084.53069	4	13021.13267	1.04
Error	350936.93555	28	12533.46198	

*Significant at the 0.10 level or better.

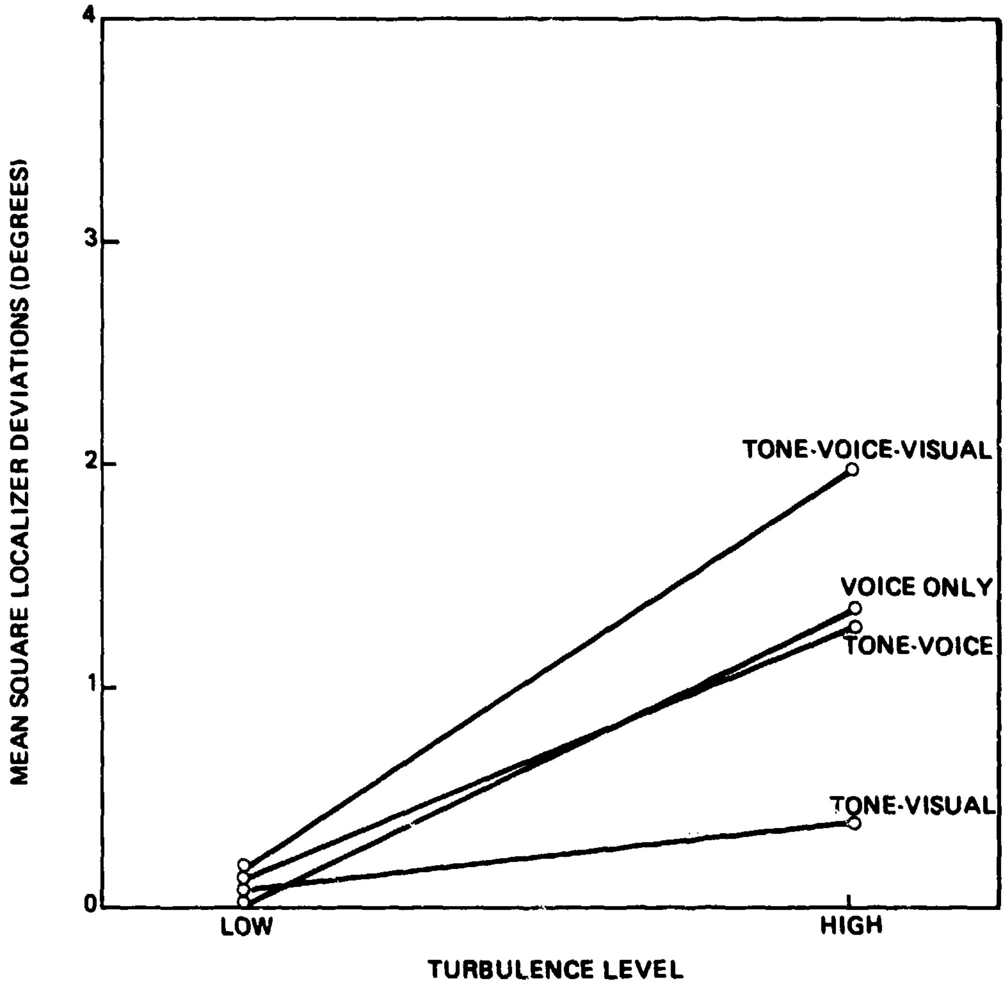


Figure 5.8.4.4-1. Mean Square Localizer Deviations as a Function of Alerting Mode and Turbulence Level: Nonconcurrent ATC Communications

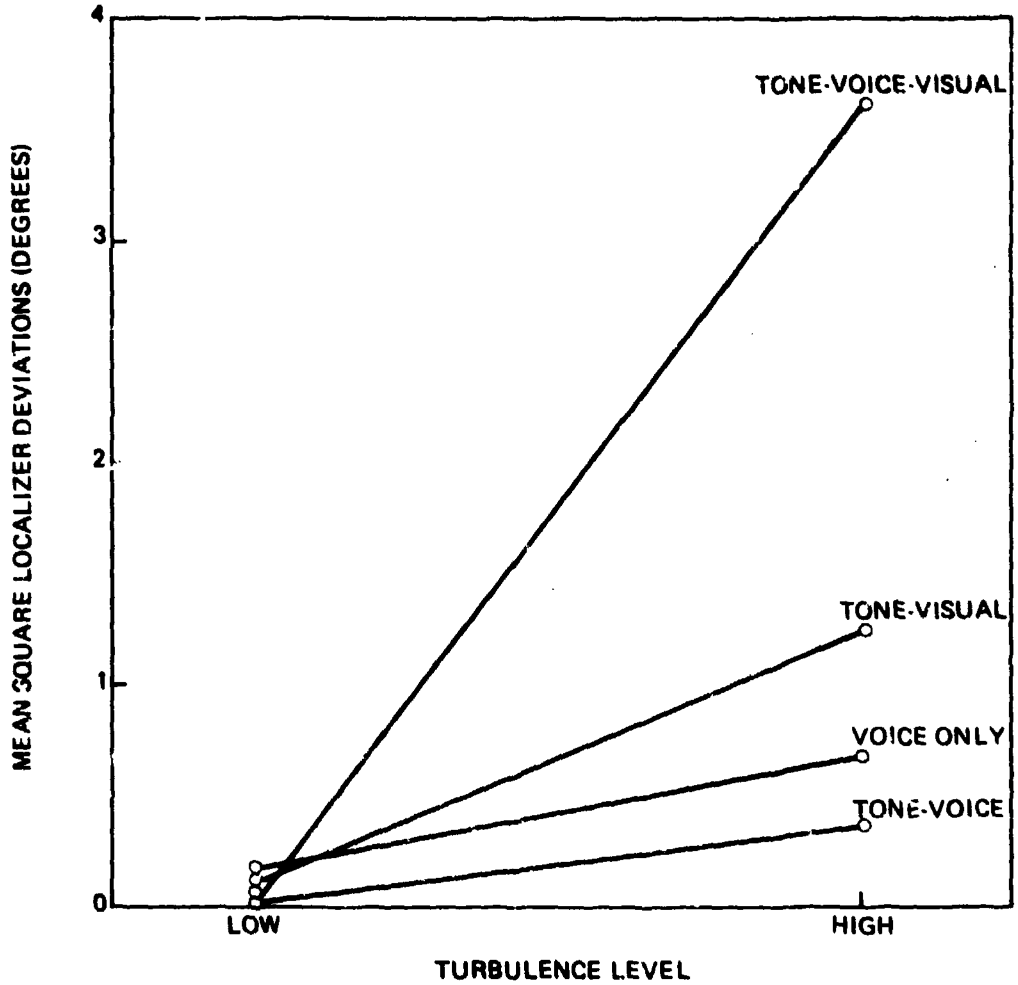


Figure 5.8.4.4-2. Mean Square Localizer Deviations as a Function of Alerting Mode and Turbulence Level: Concurrent ATC Communications

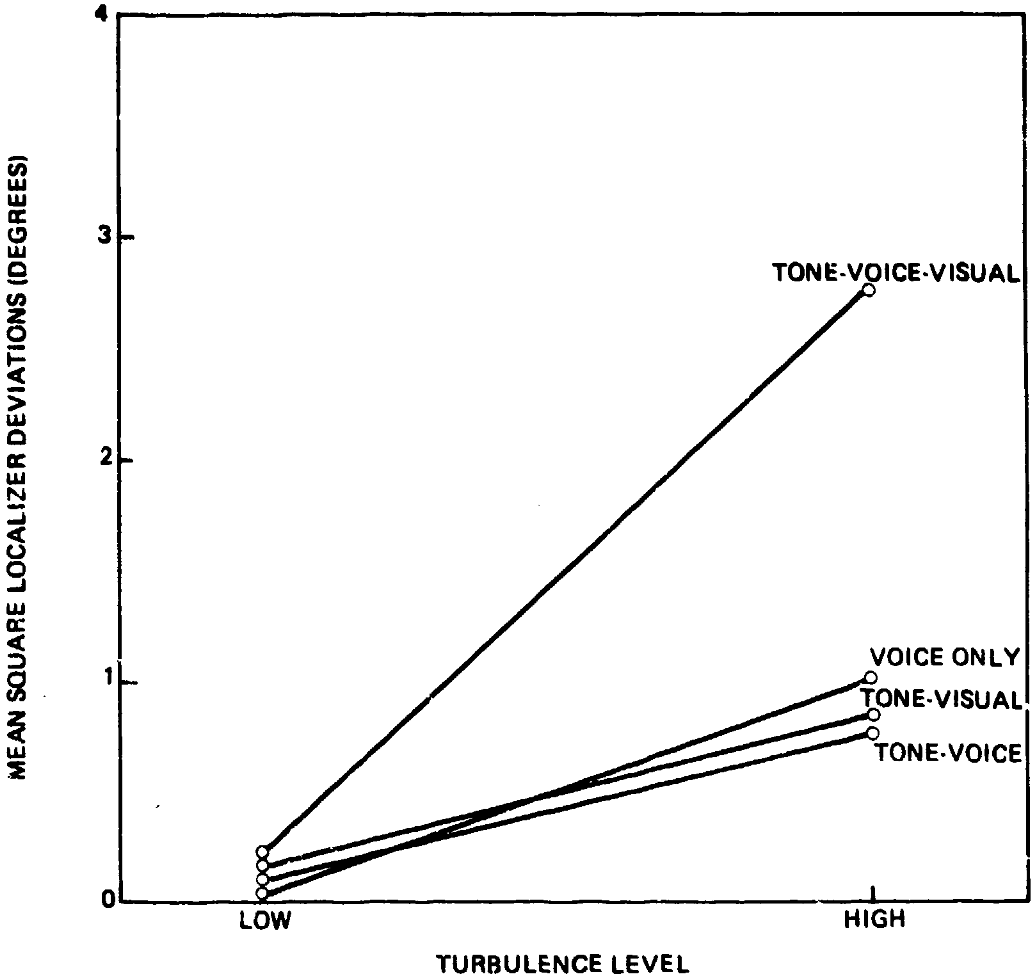


Figure 5.8.4.4-3. Mean Square Localizer Deviations as a Function of Alerting Mode and Turbulence Level: Concurrent and Noncurrent ATC Communications

mode produced considerably larger decrements in tracking task performance than the other alerting modes. This effect was statistically independent of the level of auditory task loading.

The results of the analysis of variance for roll deviation data are presented in Table 5.8.4.4-2. Data points representing the mean square roll error average across subjects are plotted for low and high turbulence levels in Figure 5.8.4.4-4. The values shown in this graph were derived by combining scores for concurrent and non-concurrent ATC communications. As anticipated, the level of turbulence had a strong effect on the pilot's ability to maintain the appropriate roll attitude. The obtained F-ratio for visual workload was significant at the $< .01$ level. The effect of turbulence level on roll deviations was consistent across alerting modes as indicated by the nearly parallel trend lines in Figure 5.8.4.4-4. Although alerting mode accounted for only a small proportion of the variance in roll error, it should be noted that roll deviation and localizer deviation measures yielded identical rank orders for alerting modes under the high turbulence condition (compare Figure 5.8.4.4-3, -4). This rank order equivalence is to be expected since rolling the aircraft provides the only method for adjusting the lateral position of the aircraft with respect to the flight director target.

Results of the analysis of variance for the third measure of tracking task accuracy are summarized in Table 5.8.4.4-3. The data plots for mean square altitude error in Figure 5.8.4.4-5 are combined as before across levels of auditory workload. The main effect for visual workload accounted for the only significant component of variance ($p < .01$). Glideslope deviation data provided no evidence of differential distraction or disruptive effects between alerting modes.

5.8.4.5 PILOT PREFERENCES

Frequency and percentage statistics describing pilot preferences for alerting mode are presented in Table 5.8.4.5-1. The tone-visual mode was chosen most frequently on the basis of attention-getting value and overall effectiveness. The combined tone-voice-visual mode was the preferred option for 3 of the 8 pilots participating in test 4. Preference judgments were distributed about equally between tone-voice and voice only alternatives.

Table 5.8.4.4-2. Test 4: Analysis of Variance Summary Table (Roll Deviation Data: Postalert Segment)

Source	Sum of squares	Degrees of freedom	Mean square	F ratio
Mean	46291.00439	1	46291.00439	86.23
Error	3757.80226	7	536.82889	
Mode	1515.24465	4	378.81116	0.82
Error	12950.65905	28	462.52354	
ATC timing	3.05257	1	3.05257	0.01
Error	3884.33953	7	554.90565	
Mode x ATC timing	1532.89889	4	383.22472	0.77
Error	13969.52805	28	498.91172	
Turbulence	31531.02633	1	31531.02633	76.09*
Error	2900.63784	7	414.37683	
Mode x turbulence	679.14255	4	169.78564	0.35
Error	13580.22666		485.00809	
ATC timing x turbulence	66.95157	1	66.95157	0.17
Error	2813.95157	7	401.95982	
Mode x ATC timing x turbulence	2176.08844	4	544.02211	1.30
Error	11681.76060	28	417.20574	

*Significant at the 0.10 level or better.

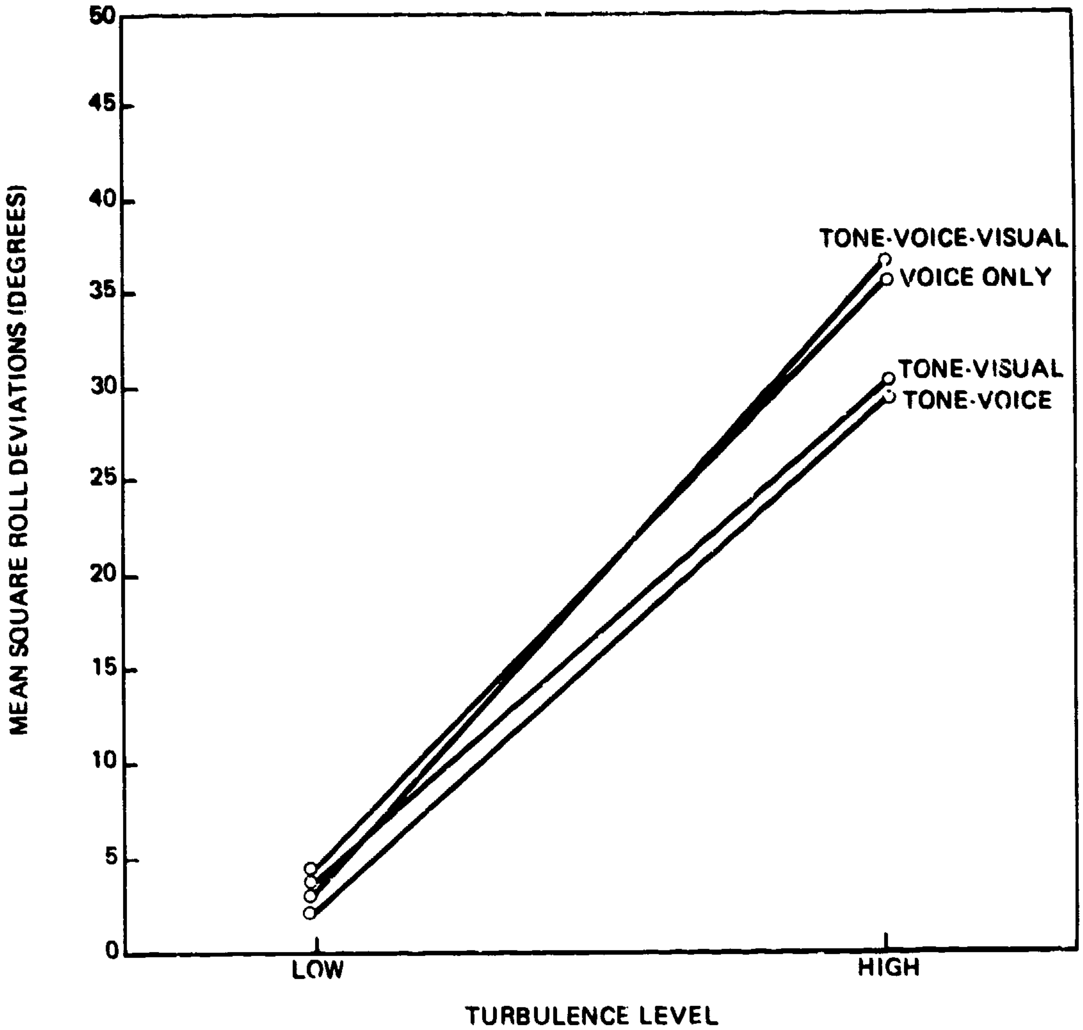


Figure 5.8.4.4-4. Mean Square Roll Deviations as a Function of Alerting Mode and Turbulence Level: Concurrent and Nonconcurrent ATC Communications

Table 5.8.4.4-3. Test 4: Analysis of Variance Summary Table (Altitude Error Data: Postalert Segment)

Source	Sum of squares	Degrees of freedom	Mean square	F ratio
Main	527936517.57677	1	527936517.57677	40.29
Error	91731414.49579	7	13104487.76511	
Mode	46172941.84770	4	11543235.46193	0.59
Error	543550655.78046	28	19412523.42073	
ATC timing	16313797.56011	1	16313797.56011	0.87
Error	130715427.80242	7	18674061.11463	
Mode x ATC timing	131631603.34718	4	32907900.83679	
Error	525919355.62677	28	18782834.12953	
Turbulence	163016045.53194	1	163016045.53194	12.23*
Error	93317425.80109	7	13331060.82873	
Mode x turbulence	52293208.00012	4	13073302.00003	0.67
Error	544954667.09754	28	19462666.68205	
ATC timing x turbulence	13735489.01550	1	13735489.01550	0.67
Error	143201687.79964	7	20457383.97138	
Mode x ATC timing x turbulence	78962903.91795	4	19740725.97949	0.89
Error	623558481.43933	28	22269945.76569	

* Significant at the 0.1 level or better.

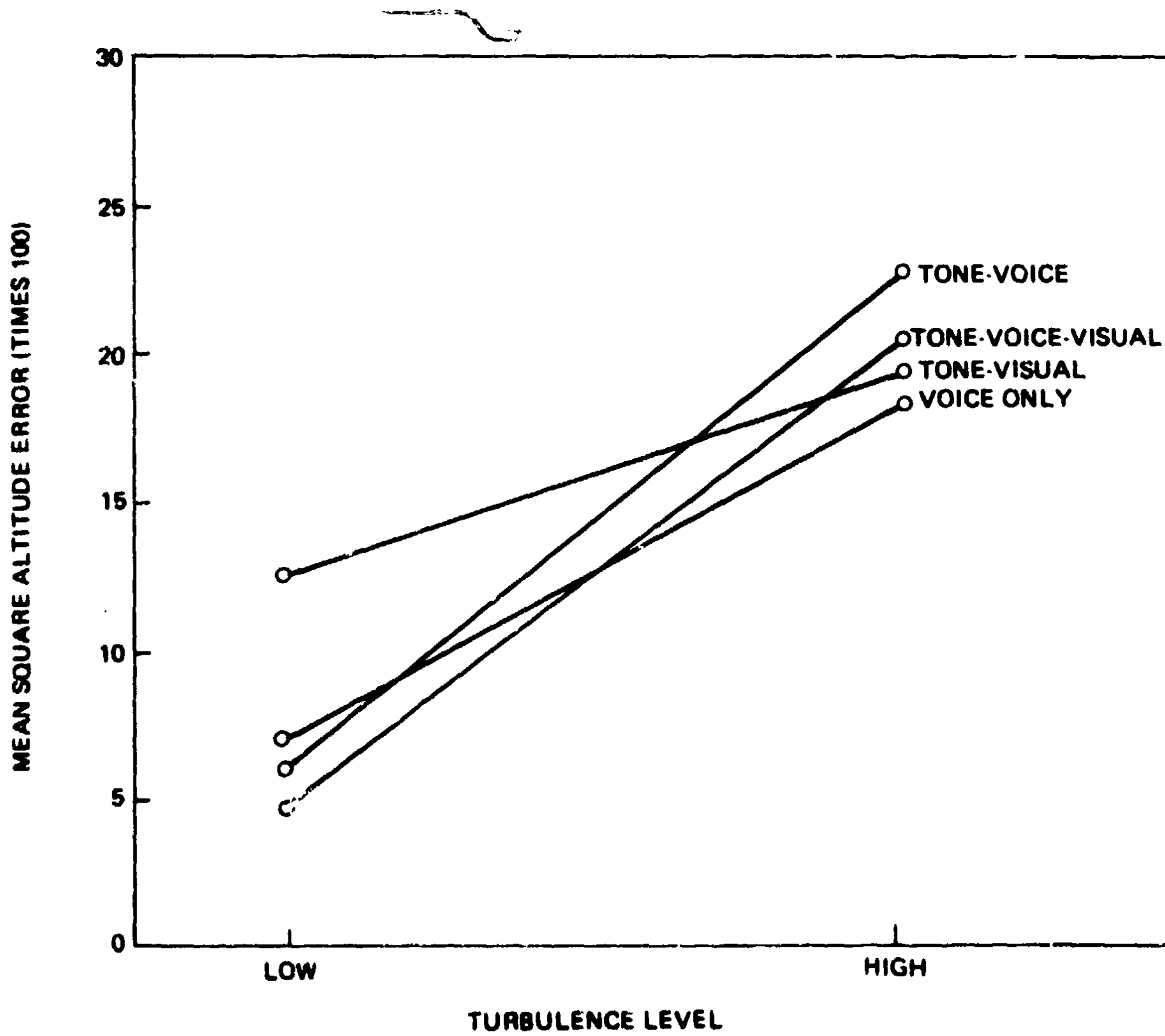


Figure 5.8.4.4-5. Mean Square Altitude Error as a Function of Alerting Mode and Turbulence Level: Combined Data for Concurrent and Nonconcurrent ATC Communications

Table 5.8.4.5-1. Pilot Preferences for Alerting Mode (n = 8)

	Tone-visual		Tone-voice		Voice only		Tone-voice-visual	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
Which alerting mode was most effective in getting your attention?	4	50.00	1	12.50	0	0.00	3	37.50
Which alerting mode would you prefer for a cockpit warning system?	4	50.00	0	0.00	1	12.5	3	37.50

5.8.5 RESULTS OF DATA ANALYSIS FROM TEST 5

5.8.5.1 VISUAL DISPLAY

The pilots' reactions to the overflow concept were quite mixed. Figures 5.8.5.1-1 through 5.8.5.1-6 present the mean responses to the questionnaire and the 95% confidence interval around these means; Table 5.8.5.1-1 presents an overall mean response for each concept.

Both groups of pilots preferred the scrolling concept; the reaction to the other concepts was highly dependent on the pilots' experience. The Boeing pilots, with more experience on Boeing aircraft, preferred the system association concept over the concept of dropping messages off the bottom of the display; the Douglas group had the reverse preference.

Both groups showed a significant preference for the unique advisory. As can be seen in Figures 5.8.5.1-7, -8 the blue advisory was rated better on all questions and there was no overlap of the confidence intervals. Table 5.8.5.1-2 presents the overall ratings of the two concepts.

When asked to comment about presenting the messages in chronological order as compared to presenting the messages category, the preference was very evenly split. 50 percent of the pilots preferred prioritization and 42.8 percent preferred chronological ordering; 7.2 percent preferred having the warnings on top and the cautions and advisories chronological (Table 5.8.5.1-3)

Fifty-three percent of the pilots who preferred the scrolling concept for overflow conditions suggested scrolling by pages rather than moving the display one line at a time. The scrolling method suggested most often was the use of multiple pushes of the recall button.

The pilots who preferred the system reversion concept noted some problems. For example, if a new alert occurs in a faulted system, the pilot may not be able to find the new alert because the display does not change. The pilots suggested that a faulted system containing only advisory alerts should be displayed differently than one which has some caution alerts.

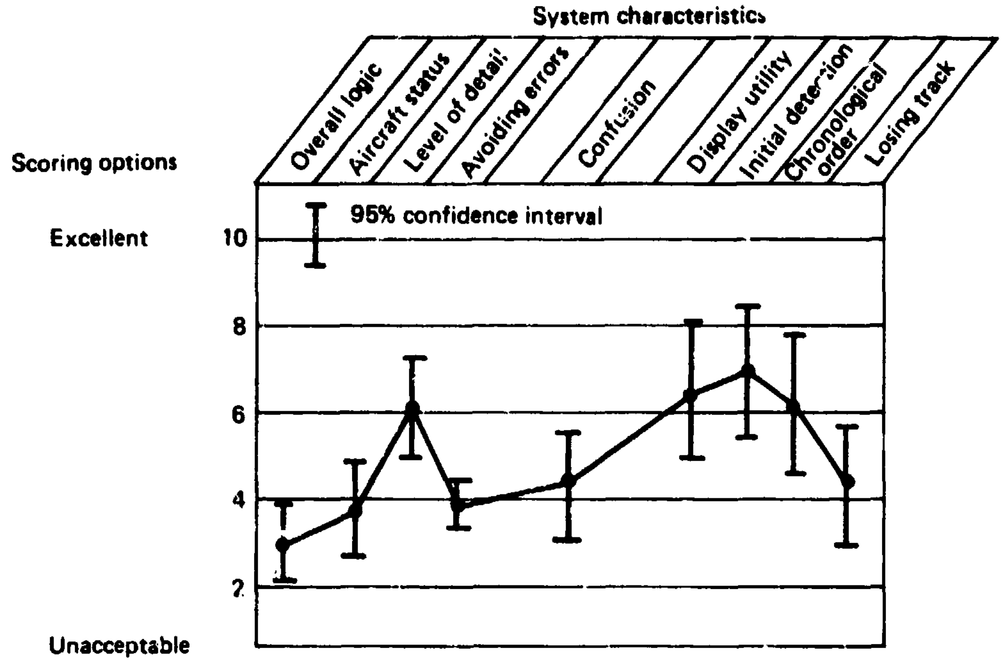


Figure 5.8.5.1-1. Boeing Pilots Mean Ratings and 95% Confidence Intervals for the Dropoff Overflow Concept

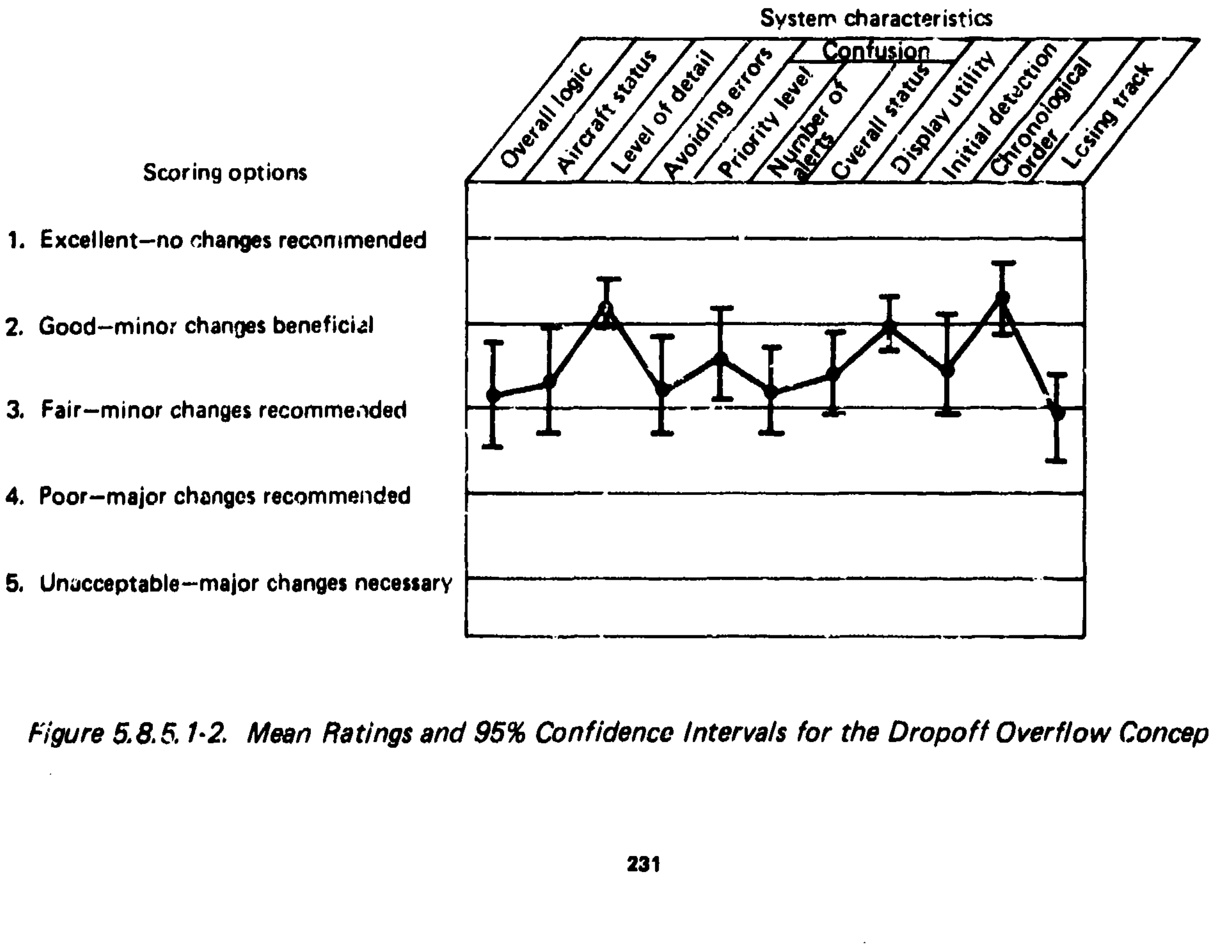


Figure 5.8.5.1-2. Mean Ratings and 95% Confidence Intervals for the Dropoff Overflow Concept

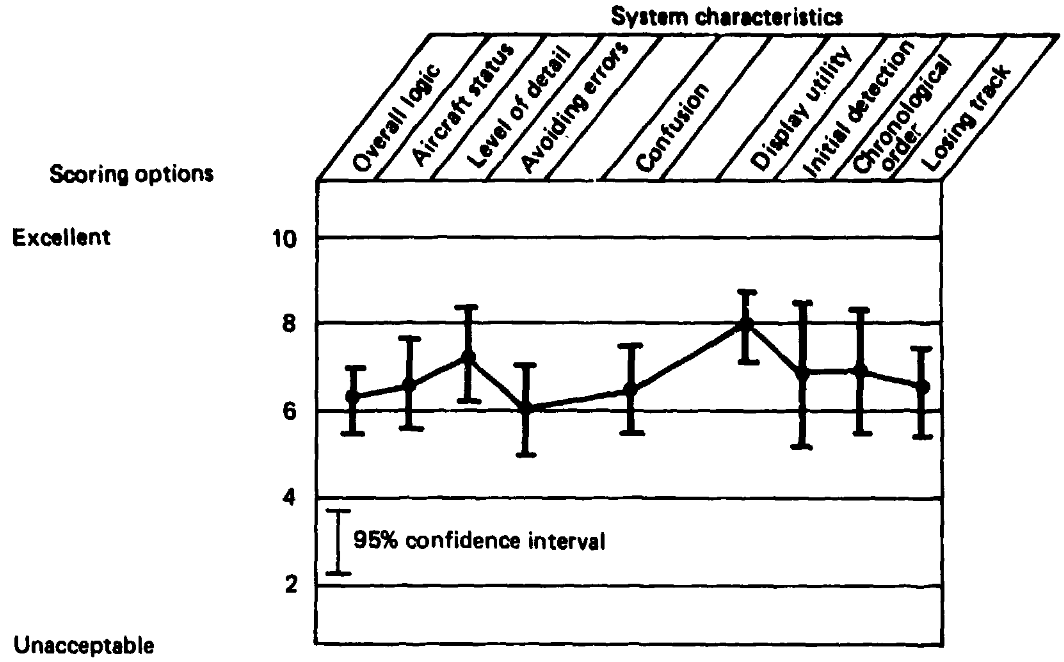


Figure 5.8.5.1-3. Boeing Pilots Mean Ratings and 95% Confidence Intervals for the Roll Overflow Concept

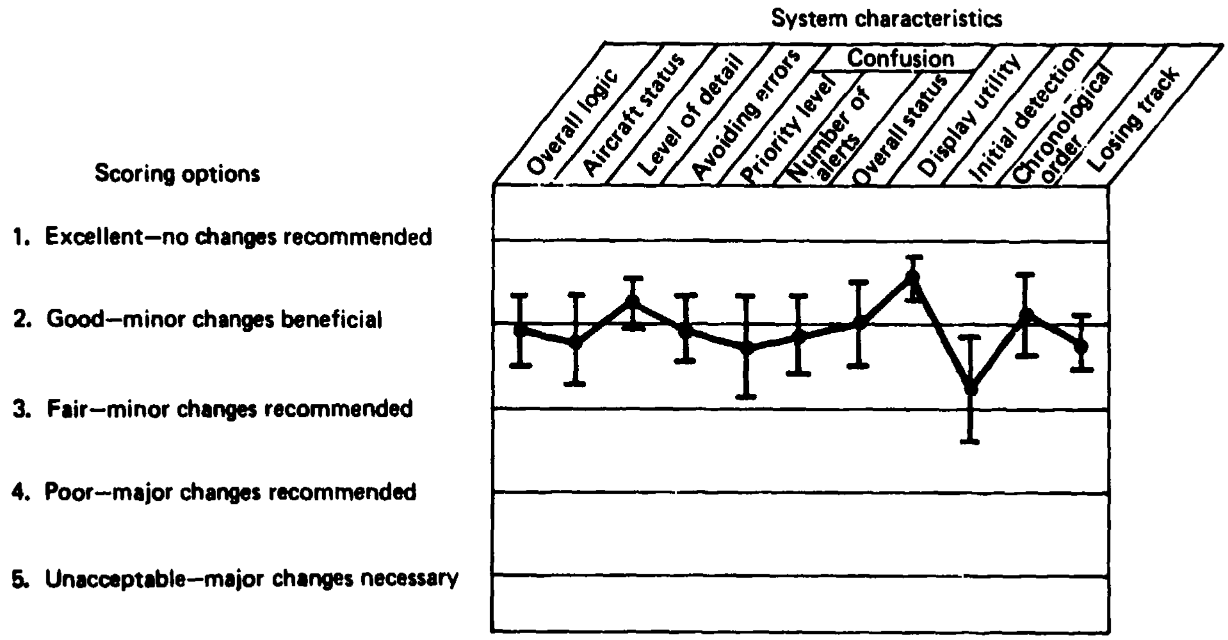


Figure 5.8.5.1-4. Douglas Pilots Mean Ratings and 95% Confidence Intervals for the Roll Overflow Concept

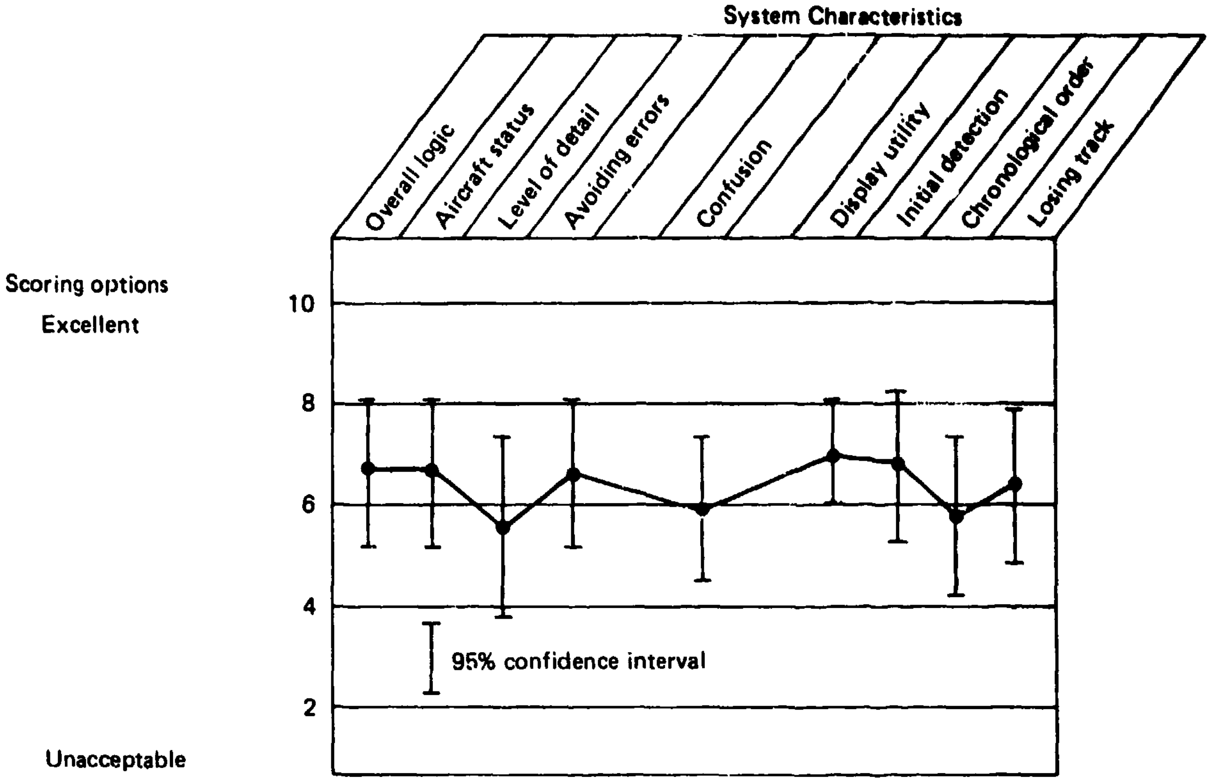


Figure 5.8.5.1-5. Boeing Pilots Mean Ratings and 95% Confidence Intervals for the System Summary Overflow Concept

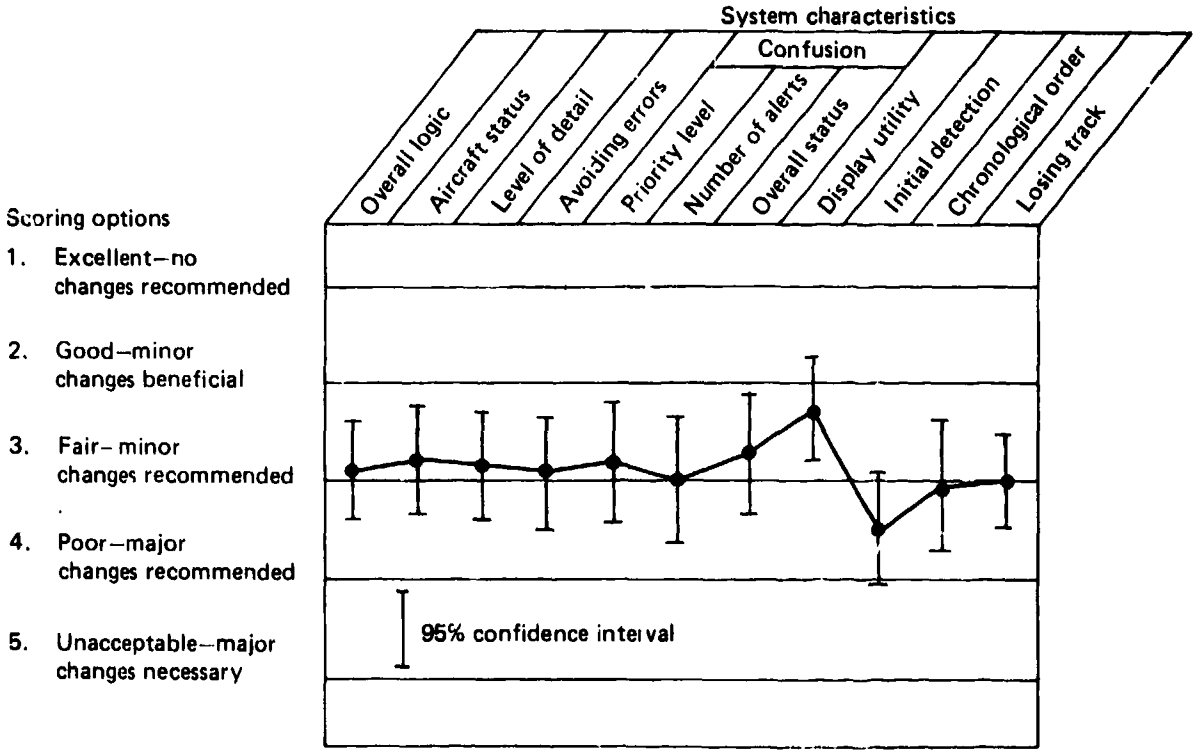


Figure 5.8.5.1-6. Douglas Pilots Mean Rating and 95% Confidence for the System Summary Overflow Concept

Table 5.8.5.1-1. Overall Summary Ratings for Overflow Logic Concepts

Overflow concept	Median rating	Mean rating	Standard deviation	95% limits
Dropoff	2.30*	2.43	0.70	2.80
	4.4**	5.05	2.14	2.07
				5.47
				4.63
Roll	1.95*	2.02	0.63	2.35
	6.5**	6.78	1.96	1.69
				7.17
				6.40
Subsystem	2.76*	2.84	0.97	3.35
	6.42**	6.37	2.65	2.33
				6.89
				5.85

*5-point scale: 1 = excellent, 5 = unacceptable

**10-point scale: 1 = unacceptable, 10 = excellent

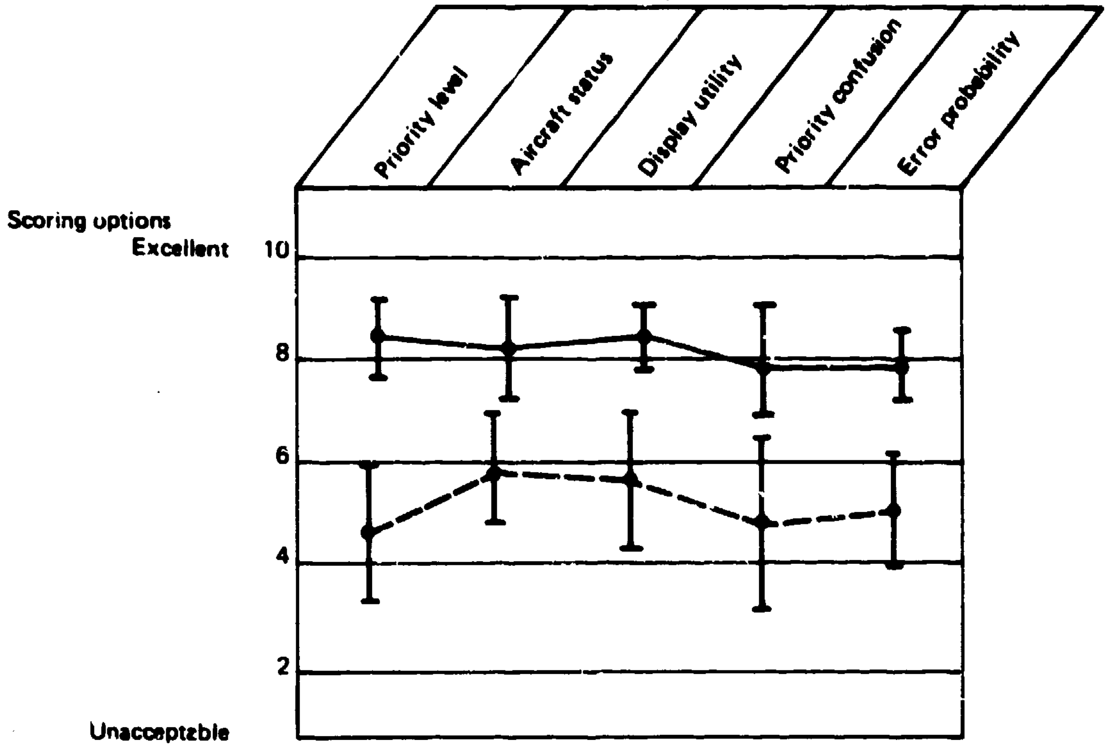


Figure 5.8.5.1-7. Boeing Pilots Mean Ratings and 95% Confidence Intervals for Blue Advisories

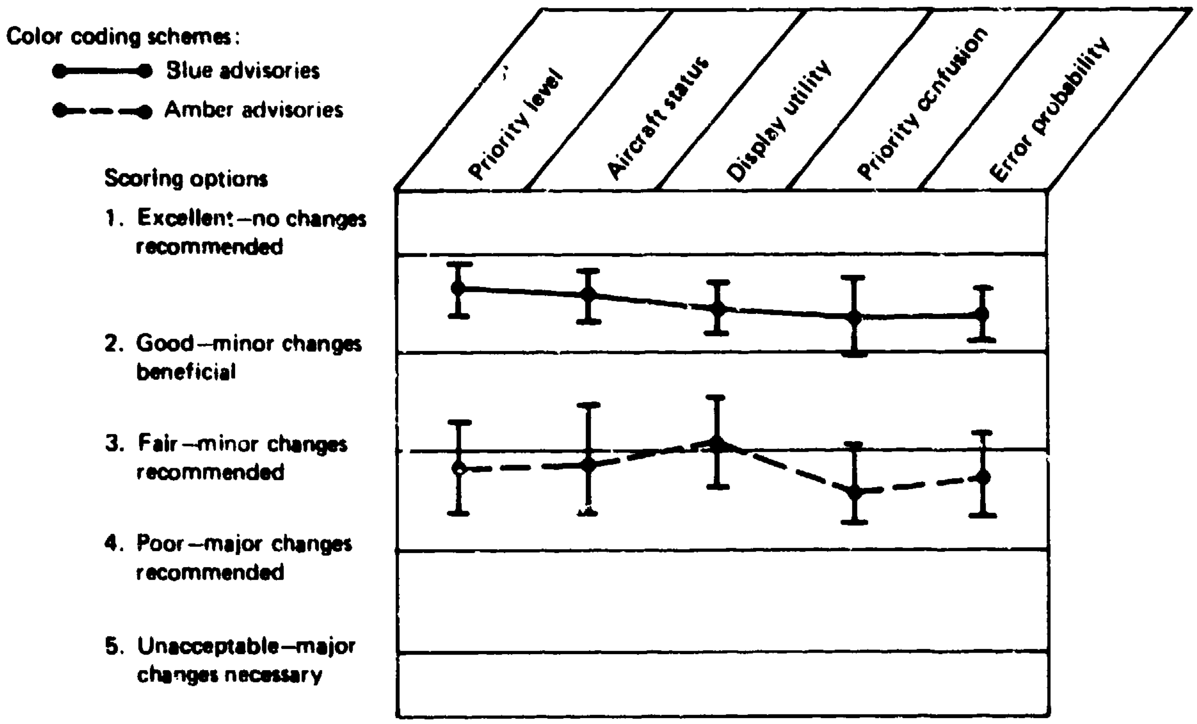


Figure 5.8.5.1-8. Douglas Pilots Mean Rating and 95% Confidence Intervals for Color-Coding Schemes

Table 5.8.5.1-2. Overall Summary Ratings for Color-Coding Schemes

Color-coding scheme	Median rating	Mean rating	Standard deviation	95% limits
Blue advisories	1.50*	1.52	0.39	1.72
				1.33
	8.17**	8.09	1.53	8.49
				7.69
Amber advisories	3.34*	3.17	0.81	3.57
				2.78
	5.5**	5.02	2.33	5.58
				4.54

*5-point scale: 1 = excellent, 5 = unacceptable

**10-point scale: 1 = unacceptable, 10 = excellent

Table 5.8.5.1-3. Pilot Preferences for Visual Display Format (28 Pilots)

Display format	Frequency	Percent
Chronological	12	42.8
Priority	14	50.0
Warnings separate	2	7.1

86% of the pilots said that there must be some indication that overflow has occurred, no matter which overflow concept is used; 54% recommended that an arrow ( ) be used to show that more information is available. Five pilots said that logic should be provided to reduce the number of alerts presented as a result of a major failure (e.g., engine failure). Finally, 82% of the pilots commented that a computer should be utilized to allow the crew to interact with the alerting system and its displays, (e.g., selective recall, selective aural/visual presentation, calling up checklists, expanding the subsystem reversion, paging or scrolling. A summary of all the comments is presented in Appendix E.

5.8.5.2 AUDITORY DISPLAY

The pilots attached meanings to all the tones presented; their individual responses were found to be highly dependent on each pilot's background. For example, 81 percent of the Douglas pilots identified the electronic bell with fire, while only 25 percent of the Boeing pilots responded accordingly. The mechanical bell elicited a fire response from 92 percent of the Boeing pilots and only 31 percent of the Douglas pilots. Boeing aircraft use a mechanical bell sound as a fire warning and Douglas uses an electronic bell; this example illustrates how a pilot's past experiences can effect alerting system designer. The final results can be seen in Table 5.8.5.2-1.

Table 5.8.5.2-2 lists the twelve sounds presented and the categories of urgency they elicited. The intermittent, wavering sound was selected most often as indicating warning, e.g., mechanical bell 100%, high wailer 92%, electronic bell 82.1% etc. The steady, insistent sound of a high horn (60.7%) and the low c-chord (57.1%) were chosen most often as cautions, and the single stroke tones such as the high and low chimes (92.8% and 78.6%) were most often considered advisory sounds.

Finally, all the pilots responded to the question which asked them whether certain specific alerts should have their own tone; this data is presented in Table 5.8.5.2-3. Only two alerts, fire and overspeed, were identified by more than 50 percent of the pilots.

Table 5.8.5.2-1. Summary of Pilot Judgment of Stereotyped Meanings for Alerting Tones

ALERTING TONE	SPECIFIC STEREOTYPED MEANINGS ASSIGNED																
	FIRE	STALL	AUTOPILOT	EVACUATION	OVERSPEED	ALTITUDE	SELCAL	TERRAIN	CABIN CALL	PRESSURE	LANDING GEAR	SMOKE-BELT	STABILIZER	STABILIZER TRIM	CONFIGURATION	TOTAL	PERCENT (28 PILOTS)
ELECTRONIC BELL	16	1								1						18	64.29
HIGH WAILER	1		5	1	8											15	53.57
CLACKER		1			18											19	67.86
LOW C CHORD	2				2	10	3									17	60.71
LOW WAILER	1		10		3			4			1					19	67.86
HIGH CHIME	1					3	1		13	1						19	67.86
LOW CHIME						5			9		1	1				16	57.14
LOW BUZZER			1		2	2							1	1		7	25.00
HIGH C CHORD						14					2			1		17	60.71
MECHANICAL BELL	16															16	57.14
LOW HORN			1			4					4					9	32.14
HIGH HORN								1	1		6			1	1	10	35.71
TOTAL	37	2	17	1	33	38	4	5	23	2	14	1	1	3	1	182	

Note: Numerical entries indicate frequency of selection

Table 5.8.5.2-2. Summary of Pilot Judgment of Selected Alerting Tone Priority Levels

Alerting tone	Priority level assigned (% of 28 pilots) Warning	Caution	Advisory
Mechanical bell	100.0	–	–
High wailer	92.0	3.5	3.5
Electronic bell	82.1	14.2	3.5
Low wailer	67.9	28.6	3.5
Clacker	60.7	35.7	3.5
Low C-chord	7.1	57.1	35.7
High horn	3.5	60.7	35.7
Low buzzer	3.5	46.4	50.0
High C-chord	–	35.7	64.3
Low chime	–	21.0	78.6
Low horn	–	25.0	75.0
High chime	–	7.1	92.8

- Criteria for assigning tones to priority levels

Warning: Emergency operational or aicraft system conditions that require immediate corrective or compensatory action by the crew.

Caution: Abnormal operational or aircraft system conditions that require immediate crew awareness and subsequent corrective or compensatory crew action.

Advisory: Operational or aircraft system conditions that require crew awareness and may require crew action.

Table 5.8.5.2-3. Summary of Pilot Opinion on Need for Dedicated Alerting Tones in Addition to Master Warning and Master Caution Sound

Condition	Alerting tone	% of 28 pilots
Fire	Bell	82
Overspeed	Clacker	50
Gear	Horn	43
Ground proximity	Voice and wailer	36
Configuration	Horn	28
Cabin pressure	Intermittent horn	25
Altitude alert	C-chord	25
Autopilot discrimination	Wailer	18
Stall	Stick shaker	18
SELCAL	Chime	11
Crew call	Chime	11

6.0 DISCUSSION AND CONCLUSIONS

6.1 VISUAL SYSTEM TESTS 1 AND 2

Tests one and two were concerned with the visual portion of the crew alerting system and tested detection, response and accuracy of response to alerts at three levels of urgency. This section of the report discusses the major topics of concern for tests 1 and 2 including the master visual attenson, the location of the information display, the format of the material on the display, and the effect of pilot workload on crew alerting.

6.1.1 MASTER VISUAL ATTENSON

The primary function of the visual attenson is to attract the pilot's attention; it is therefore expected to see the effects of the attenson on detection times and missed alerts. The results of both experiments demonstrated that the mean detection times, for those conditions where there was no master visual alert, were significantly slower than with the alert. In the first experiment, the alerts were presented without a master attenson and had a mean detection time 63 percent slower than for alerts presented with a master (3.54 sec vs. 2.17 sec). The same finding occurred in the second experiment; mean detection times were 87% slower (2.77 sec vs. 1.48 sec).

A flashing box around the message on the information display was used as an attention-getter in the second experiment. Overall, this was not as effective as the master light on the glareshield; mean detection times were 53% slower with the flashing box (2.26 sec vs. 1.48 sec). The location of the attenson, however, had a large effect on its detection. The master alert light was always in the pilot's primary field of view; the flashing box, on the other hand, had to be located with the display and therefore, in some of the test conditions, was located outside the primary field of view. Figures 4.8.2.1-3, -4, and -5 show that the mean detection time for warnings with a flashing box as the attention-getter was significantly slower when it was on the center screen (3.06 sec) than when it was in front of the pilot (1.6 sec); the value for the condition with the attenson and the flashing box was 1.2 sec.

The format of the master alert does not seem to have an effect on the mean detection time. The large surface area of the single master was not detected significantly faster than the smaller alert in the dual system. The flashing alert also did not produce a significantly shorter mean detection time than the single or dual master system. These results were augmented by the comments of some of the pilots who felt that both the single master and the flashing masters were distracting. As one pilot put it: "The flashing light is very attention-getting but it is more likely to provoke a 'knee jerk' reaction (perhaps the wrong one) by the pilot".

Another side effect of the master visual attenson demonstrated by both experiments was its ability to change the pilot's scan pattern. When the master attenson contained no information (single system) or when it was absent, the detection times for advisories was reduced; this apparently reveals that the pilots were scanning the display more frequently. However, when they could depend on the master to provide the pertinent information, they scanned the display less often and the detection times for the advisories increased.

The number of missed alerts were highly dependent on the presence of an attention-getting device. In experiment one, 80 percent of all the alerts missed (24 out of 30) occurred when there was no master attenson. Similar results were obtained in experiment two; 88 percent of the missed alerts (65 out of 74) occurred when no attention-getter was presented. With respect to missed alerts, the type of attention-getter used did not significantly affect the results.

It was not expected that the attention-getters would affect response times, other than that part of the response time consisting of detection. This, however, was not the case; the pilot used the information present in the master alert to make a decision about his response. It was consistently demonstrated that if the pilot could determine whether the alert was a caution or a warning without looking at the information display, he would respond more quickly to the warnings than to the cautions. This occurred even though the mean detection times for the cautions and warnings were not significantly different.

The pilots preferred the master visual attenson to the flashing box and felt that a combination of the two would be helpful in finding the most recent alert. When asked about the features they most liked about the system, 71 percent of the pilots listed the master attenson and 57 percent the flashing box.

Finally, the type of attention getting devices used did not have a measurable effect on the performance of the flight task; the same was true for the format of the master attenson. Some of the pilots felt, however, that the flashing light did have an effect.

6.1.2 DISPLAY FORMAT

Message format had no measurable effect on the subjects' response to the alerts as far as the main effect was concerned. However, the interaction between format and target type revealed that cautions had a significantly shorter mean response time when they appeared at the top of the display screen than when they appeared at the top of their own category (5.48 sec vs. 6.75 sec). This result indicates that the pilots took more time to find the cautions when they were presented in the middle of a group of alerts. One way to alleviate this problem would be to somehow show the pilot which is the most recent alert. Test 2 had such a condition i.e., a flashing box was used to indicate the latest caution along with the master visual attenson. Test 2 duplicated the condition from test 1 where both cautions and warnings were presented with a master attenson and appeared on the same screen by category with no flashing box. The mean response time for this condition was comparable for both experiments (6.75 sec vs. 7.12 sec). Adding the flashing box, however, resulted in a significantly faster mean response time (5.56 sec); in fact, this response time is comparable to the condition in test 1 which had the cautions appearing on the first line (5.48 sec). It can therefore be concluded that, some indication should be used to designate the most recent alert if that alert requires immediate attention and if it has the potential for appearing on any line except the top line.

The format of the messages on the display did not have a measurable effect on the number of missed alerts. The number of advisories missed with messages separated by urgency levels was, however, slightly higher than for the other two concepts.

The pilots were divided on their preferences; 50 percent preferred the categorization of alerts by urgency and 36 percent by chronological order with the latest message entering on top. The second concept, it was felt, would aid in finding the most recent alert; however, categorization was preferred for assessing aircraft status because in a display overflow situation, the least important messages would leave the screen first.

There was no significant effect on the flight task that could be attributed to the information display message format or to the interaction of message format with attenson format or to the flashing box.

6.1.3 DISPLAY LOCATION

It has long been recognized that the most ideal location for alerting displays is within the pilot's primary field of view. The question is, how much information should be put in this area? Is it adequate to just get the pilot's attention and inform him of the urgency of the alert, or should more information be provided for high priority alerts?

If a master visual attenson is used, the location of the information display has no measurable effect on either alert detection or response. However, if no master attenson is used, but a flashing box or some other attention-getter is used on the display itself, location does have an effect. For those conditions with a flashing box and no master attenson, the mean detection and response times for warnings were significantly shorter when the display was located within the pilot's primary field of view. When both warnings and cautions were presented in the primary field of view with a flashing box and no master attenson, there was no measurable difference in the mean detection times; the mean response times were significantly different as they were with the master attenson.

There were no measurable differences in the number of missed alerts that could be attributed to the location of the display; however, the number of advisories missed was slightly higher when all the alerts were on the center display. This may be because when all alerts are presented on the center screen, the advisories appear between two other messages. This hypothesis is supported by the fact that the least advisory misses occurred when advisories alone appeared on the display and the most recent alert was presented on the top line.

The pilots showed a clear preference for presenting all alert messages on a center display. They felt that it was confusing to split up the information; the master alert should tell them to look only one place. They also said, however, that it was harder to spot the advisories when other alerts were on the display on top of them.

As with the other alert variables, there was no demonstrable effect of display location on the pilot's flight task performance.

6.2 AURAL SYSTEM TEST 3

6.2.1 CONFUSION - MASKING EFFECTS

Test 3 results indicate that there is a significant potential for mutual masking between synthetic speech warning messages and other voice communications within the cockpit. These effects occurred even though the auditory alerting system was designed to optimize the pilot's ability to discriminate and selectively attend to the two sources of auditory information. Performance on the ATC recognition task was substantially degraded on those trials when tower advisory messages were presented concurrently with voice alerts. The maximum frequency of errors for the non-concurrent messages was 12.5 percent while error rates for concurrent messages ranged from 75 to 94 percent.

Masking effects between competing speech sources were also evident in the response time data. The time required to detect, identify and acknowledge an alert was increased considerably when messages were presented simultaneously; the F-ratio for the Auditory Workload main effect was significant at the .05 level. Figure 5.8.3.1-3 shows that increments in mean response time resulting from concurrent messages varied from one to two seconds, depending on alert format.

6.2.2 CONTROLLER VOICE QUALITY

The quality of the air traffic controller voice had no measurable effect on the test subject's ability to discriminate between alert and ATC messages. The anticipated decrement in performance due to simultaneous onset of two female voices was not evident in the test data. The obtained F-ratio associated with the ATC voice quality main effect response time did not approach statistical significance; Chi-square tests comparing ATC recognition error marginal totals for the male and female controller voices yielded no reliable differences. The error distribution graphs in Figures 5.8.3.3-3 and -4 actually suggest a slight improvement in ATC recognition accuracy for the female controller; these data indicate that the female controller voice may have been slightly more intelligible than the male voice. The voice model

used in programming the central aural warning system was equally intelligible with respect to the two controller voices selected for this study. Results of the error analysis indicate that the proportion of relatively serious types of errors was independent of the nature of the controller voice.

These results are not surprising when considered in the context of existing data on human voice intelligibility; tests of speech intelligibility tend to yield large individual differences, regardless of the sex or register of the speaker. It seems probable that the articulation characteristics and speech rate of the individual controller, and the quality of the transmission system are the most important factors, in determining the accuracy of ATC recognition.

6.2.3 ALERTING TONE

The presence of a precursor alerting tone did not enhance the attention-getting value of the voice alert in terms of time required to initiate corrective action. Inspection of the response times plotted in Figures 5.8.3.1-1, -2 and -3 show somewhat longer mean response times for messages preceded by a tone. This effect may be due to the delayed onset of the critical elements of the voice message when an alerting tone is used. These observations are consistent with the findings of a previous study conducted by NASA (Simpson and Williams, 1978). Results of the NASA study showed significantly longer response times for synthesized speech warnings when they were preceded by an alerting tone of 500 msec.

The presence of an alerting tone had no effect on the frequency of ATC recognition errors or the distribution of error types; the total error frequencies obtained for concurrent messages were identical for the tone-voice and voice only alerting modes. The data provide no direct evidence of increased confusion, distraction, or masking when an alerting tone is present. Although the alerting tone provided no measurable benefits in terms of objective performance data, the subjective data summarized in Table 5.8.3.5-1 indicate that the tone-voice option was preferred by an overwhelming majority of the pilots tested. This strong preference may be due in part to their familiarity with alerting tones as attention getting devices and their relatively limited exposure to the voice-only concept.

Care should be exercised in generalizing beyond the specific alerting tones and sound levels employed in this experiment. Sound levels and tone qualities were selected based on preliminary research to insure adequate intelligibility and minimum annoyance. Alerting tones may have important attention-getting or noise-penetrating characteristics in severely degraded auditory environments. It should also be noted that realistic attentional demands and levels of vigilance are very difficult to achieve in a simulated cockpit setting. In an operational cockpit environment, the onset of a warning or caution level aural alert is a relatively infrequent event. Under these conditions, the attention-demanding characteristics of a precursor tone might be more readily apparent. The data indicate that some existing guidelines and specifications requiring attention-getting tones for all voice alert messages may be unwarranted.

The additional value of tones as a source of alert priority level information should not be overlooked; individual tones corresponding to distinct levels of urgency might be used to facilitate the crew's decision making process and expedite corrective action under emergency conditions. It appears, however, that the effectiveness of tones as a supplement to synthetic voice alerts depends largely on the particular application and cockpit environment.

6.2.4 VOICE MESSAGE STRUCTURE

Improvements in message intelligibility and response time associated with increased linguistic context were reported by Hart and Simpson (1976). These findings are supported to some extent by the results of the present study. Although no significant response time main effects were recorded as a function of voice message format, the format by auditory workload interaction suggests that the sentence format may offer some advantages when messages are concurrent. Some of the gains in intelligibility due to purely redundant language may be offset however because of the increased time required to annunciate the essential message elements.

Simpson and Williams (1978) pointed out that the complete sentence message format may offer additional advantages by providing more specific information about the nature and source of the problem. This viewpoint was corroborated

somewhat in the present study by several test subjects who felt that complete statements such as "THE AUTO SPOILERS ARE NOT ARMED" would be more informative in an operational context than brief phrases like "AUTO SPOILERS".

Inspection of the ATC error contingency tables indicates that the distribution of error types was essentially constant across alert message formats. The ATC recognition accuracy data do not support the hypothesis that longer voice alert messages increase the probability of interference with communications.

The findings suggest that pilot information requirements and the nature of the corrective action should be the determining factors in structuring voice messages. Assuming that the flight crew is familiar with ATC terminology and the alert message set, linguistic considerations appear to be relatively unimportant.

6.3 VISUAL AND AURAL SYSTEM-TEST 4

6.3.1 VISUAL WORKLOAD

The tracking task error measures provided objective indices of the pilot's visual workload. Introduction of high levels of turbulence into the wind profile resulted in significant increases in deviations from the flight director target recorded during post-alert segments. The large proportion of variance in these measures attributed to turbulence level indicated that visual workload was manipulated through a considerable range.

Pilots were able to maintain a relatively stable level of performance on the other assigned tasks across low and high visual workload conditions; the results do no support the hypothesis that response times to alert messages would be significantly increased under high levels of visual workload. Performance on the ATC recognition task was also resistant to degradation as a function of visual task loading; the attentional demands of the two-axis tracking task apparently did not interfere with the pilot's ability to detect and interpret ATC or alert messages.

6.3.2 AUDITORY WORKLOAD

With respect to potential confusion and masking effects, the results of test 4 conform closely to the findings of test 3; the difficulty of discriminating between competing speech sources was demonstrated dramatically. The concurrent condition was associated with longer alert response times and substantially higher error rates on the ATC recognition task; this happened even though these data were obtained in an environment that was optimized for selective attention.

6.3.3 ALERTING MODE

In general, the alerting mode used did not affect the time required to identify and acknowledge alert messages significantly. Inspection of the response time trends in Figure 5.8.4.1-1 suggests that alerts presented by means of the tone-visual mode are subject to minimal confusion and masking

effects as a result of concurrent ATC communications. Similar relationships were obtained for ATC error frequencies as a function of alerting mode and ATC message timing. The ATC recognition accuracy data plotted in Figure 5.8.4.2-1 indicates that tone-visual alerts caused negligible increases in error rates under the concurrent message condition. The probability of mutual interference between alert messages and voice communications is reduced significantly when the tone-visual alerting mode is employed. On the basis of observed response times and total error frequencies, the tone-visual mode seems to be most effective overall for the concurrent messages. Based on post-test pilot judgments, the tone-visual mode was also determined to be the preferred option in terms of attention getting quality and overall effectiveness.

The results of the ATC recognition error analysis place some qualifications on these findings. While the overall number of errors was lowest for the tone-visual alerting mode, the proportion of relatively serious types of errors was larger but not significant. Combined omission and readback errors (4) represented 80 percent of the total number of errors for the tone-visual mode (5); the corresponding error rates for the voice alert modes ranged from 9 (1 out of 11) to 22 (2 out of 9) percent.

The potentially serious nature of "readback" errors can be illustrated by examining specific events which were observed during the course of experimental trials. As discussed previously, a "readback" error involves the misinterpretation of an ATC instruction. On two occasions involving different test subjects, a traffic advisory message ("Traffic to your right is for runway 25 right") was interpreted as a request to change to an alternate runway ("Change to runway 25 right"). Considering the relative position of the two aircraft in question, if ATC did not correct the readback, the potentially catastrophic consequences of executing such a maneuver during an actual approach are obvious. In both instances, these readback errors occurred when the tone-visual alerting mode was operational.

The variation in error distributions across alerting modes is difficult to interpret within the context of the other test results. A possible explanation for this effect was proposed by one of the test subjects during a

debriefing session. Due to the short duration of the alerting tone, the tone-visual mode provides an opportunity to hear a significant portion of the ATC communications, even when alert message onset is simultaneous. Although test subjects were instructed to emphasize accuracy in their responses, there may be some tendency to attempt a verbal readback when the probability of misinterpretation is perceived to be low. This is especially true of simulation studies where the consequence of the error might also be perceived as low and could artificially inflate the number of interpretation errors in the simulator.

Some evidence was obtained to indicate that the onset of alert messages may have disruptive effects on accuracy of aircraft control. Although all three measures of tracking task accuracy showed significant decrements as a function of turbulence level, localizer deviations proved to be the most sensitive measure of differential distracting effects between alerting modes; the alerting mode main effect and mode by visual workload interactions accounted for significant proportions of variance in post-alert localizer error. The graphic representations in Figures 5.8.4.4-1, -2 and -3 show clearly that the combined tone-voice-visual mode was associated with large increases in lateral deviations from the flight director command symbol. This distracting effect was noted for both concurrent and non-concurrent messages.

The simultaneous onset of a tone-voice message, master light and an alphanumeric readout in the primary field of view seemed to produce a stimulus overload condition. Direct observations during experimental trials and a post-test review of the video tapes revealed some evidence of confusion on the part of test subjects and a tendency to cross-check several sources of information prior to responding. At high levels of tracking task difficulty, the pilots apparently found it necessary to sacrifice some degree of accuracy in aircraft control to identify and respond to the alert message. This result is particularly noteworthy since the redundant information presented in the tone-voice-visual mode was expected to aid pilot performance under high workload conditions.

It should be pointed out that the distracting effects observed in test 4 refer only to localizer tracking performance; glideslope error data showed no

indication of excessive deviations during the post-alert segment for the tone-voice-visual mode. When faced with an overload condition, the pilots evidently tended to devote their attention to the control dimension most critical to safety of flight.

In summary, the results confirm the hypothesis that no single alerting mode is most effective under all combinations of environmental conditions. The test data indicate that the relative effectiveness of alerting modes is dependent, to some extent, on the levels of visual and auditory task loading. In light of these findings, consideration should be given to the development of a system that would automatically change modes to compensate for alterations in attentional demands imposed by the flight environment. An adaptive system of this type would employ computer logic to modify or selectively inhibit components of the alert message for a given flight segment. For example, the voice component of an alert message might be inhibited during phases of flight such as final approach where high levels of auditory workload and frequent verbal communications are anticipated. The implementation of such a concept would provide the system designer with a degree of flexibility in attempting to maximize the attention-getting value of the alert message while minimizing disruptive effects on other essential flight crew tasks. Such a phase-adaptive alerting system may also provide a mechanism for reducing overall flight crew workload under emergency or abnormal conditions.

6.4 SUBJECTIVE EVALUATION - TEST 5

6.4.1 VISUAL EVALUATION

Pilot opinions on the alternative concepts for handling situations when there were more messages than could be displayed (system overflow) was quite mixed. The ability to see all the alert messages (rolling them up and down) was preferred over the concept in which the caution and advisories revert to a system designation. This is understandable due to the increased amount of information available. The data indicate that moving the alerts one at a time is cumbersome and increases workload possibly at a time when the pilot is already under a workload strain, (e.g., during multiple failures or major system failure). It is recommended that the system interaction be done using paging of the messages. The data also indicate the need for some indication when the system has overflowed.

The pilots expressed a significant preference for a format which grouped the alerts into unique and easily distinguishable categories. For test 5 the uniqueness was created by color coding the alerts by urgency level: red/warning, amber/caution and blue/advisory. The alternative to this code was to use color to separate the two highest level alerts (red/warning and amber/caution) and to use position (a one space indent) to distinguish between the cautions and advisories. The pilots felt the code should be consistent and the categories should be unique. They consistently rated the unique advisories (blue) higher than the position-coded advisories. From these findings it is concluded that unique categories should be a system design recommendation.

6.4.2 AUDITORY EVALUATION

All of the twelve tones presented were recognized or at least identified as having some specific meaning to a portion of the pilots. A pilot's past experience plays an extremely important role in his interpretation of the meaning of a sound. If the new system uses sounds for the master aural alerts that have been used previously, the potential exists for a pilot under stress to misinterpret an aural alert. Therefore, the sounds selected for the master

alerts in the candidate concept systems for the Phase III effort will not be sounds that have been used previously in a transport aircraft.

Selection of the sounds to be used in the candidate systems was based in a large part on the data obtained in test 5. The pilots categorized each tone with respect to the urgency they felt the sound elicited. They felt that wavering, intermittent sounds had the highest urgency value, while steady insistent sounds indicated intermediate urgency and single stroke sounds the lowest. These results indicate the sound characteristics that should be used for each of the master aural alerts. Taking advantage of the pilots' preconceived notions as to how alerts should sound could reduce the potential for confusion.

7.0 CANDIDATE SYSTEM DEVELOPMENT

The final step of Phase I was the development of candidate crew alerting system concepts; these were defined as a specific set of system characteristics which met the objectives of a advanced crew alerting system. In order to perform this task it was necessary to review the crew alerting data base and utilize the relevent information. The program ground rules and system assumptions (Section 2) provided a solid foundation. The literature (Section 3), test results (Section 5), and pilot's subjective input (Section 5), were used to identify the actual system characteristics.

The candidate systems, will be evaluated in Phase III of the program. The candidate systems will be implemented in simulation hardware and validated by comparing them to a conventional crew alerting system. Therefore, it was also necessary to define a baseline conventional crew alerting system to be used in these comparisons.

In the process of identifying the characteristics of the candidate system a number of questions arose affecting implementations. These questions will be answered before system validation so that the data can be utilized during testing. It was also recognized that some aircraft conditions or situation may require that a crew action be taken in an extremely short period of time. These alerts were defined as "time critical" warnings and a need was identified to determine appropriate presentation media and format for this type of alert.

7.1 ALERTING SYSTEM DESIGN OBJECTIVES

A number of system design objectives were used in identifying the system components, design characteristics and display logic. A major objective was to increase the efficiency of the alerting process as opposed to conventional systems which require the pilot to search out information located in numerous places throughout the flight deck. The alerting system should minimize the time for the flight crew to detect, assess and respond to alerts. It should reduce demands on crew information processing and memorization capabilities. A quiet dark cockpit should be maintained when all systems are operating

normally. Distracting effects of other flight crew tasks such as aircraft control or ATC communications should be minimized. The system should be flexible enough to permit growth without necessitating additional discrete annunciations. A final objective is to provide a system which can become a standard not only across airframe manufacturers but also across aircraft types and commercial operators.

7.2 CANDIDATE ALERTING SYSTEM CONCEPTS

The major objectives of Phase I of this program were to identify the primary components of alerting systems, obtain data for developing advanced alerting system concepts, and to define alerting system categories.

Four primary components were identified:

- Master visual attenson(s)
- Master aural attenson(s)
- Visual information display
- Voice information system

The data collected during Phase I was used to identify elements for each of these components to facilitate crew detection and response to alerts in an all electronic cockpit; these elements were then combined to derive two alternative system concepts. These concepts were then developed to cover all of the alerting requirements for a four level alerting system, (warning, caution, advisory and information). Table 7.2-1 defines these levels and illustrates the alerting system elements required to mechanize them.

Two alternative system concepts were identified. The first concept was defined by identifying all of the elements of the four components and then selecting characteristics best suited for advanced cockpits. In several instances it was deemed necessary to perform small side studies to resolve questions which arose in developing the concepts. The second concept was developed because insufficient data was available for deciding between component characteristics, and because system operation would be affected dramatically due to the option selected.

Table 7.2-1. Guidelines for Standardizing Alerting Functions and Methods

Condition	Criteria	Alert system characteristics		
		Visual	Aural	Tactile
Warning	Emergency operational or aircraft system conditions that require immediate corrective or compensatory crew action	Centrally located alphanumeric readout (red)	Attention-getting tone plus voice*	Stick shaker (if required)
Caution	Abnormal operational or aircraft system conditions that require immediate crew awareness and require prompt corrective or compensatory crew action	Centrally located alphanumeric readout (amber/yellow)	Attention-getting tone plus voice*	None
Advisory	Operational or aircraft system conditions that require crew awareness and may require crew action	Centrally located alphanumeric readout (unique color)	Attention-getting tone	None
Information	Operational or aircraft system conditions that require cockpit indications, but not necessarily as part of the integrated warning system	Discrete lights (green and white)	None	None

*Voice is optional.

The two candidate system concepts differed only in the treatment of the voice information system. The components of these alternate concepts and the side studies required to resolve component selections are described in the following paragraphs.

7.2.1 MASTER VISUAL ATTENSON

The elements of the master visual attenson for the two alternative concepts (A & B) are shown in Table 7.2.1-1. Since the two highest urgency levels (warnings and cautions) require immediate crew awareness, master visual attensons located within 15° of the pilot's centerline of vision are required. The function of these attensons is to "alert" the crew of the occurrence to a high priority event, and to guide crew action by indicating the level of the alert, (warning or caution). Due to the distracting effects of flashing lights in a cockpit, and since the attensons are located within the pilot's primary field of view, steady-state attensons were selected. The attensons will be 10% brighter than surrounding lights, and will vary between 15 to 150 ft. L., depending upon ambient light conditions.

The attensons will subtend 1° of visual angle and will illuminate red for warnings, and yellow for cautions. They will be cancellable manually or will clear automatically when faults are corrected.

Two separate side studies will be conducted in Phase III to optimize the effectiveness of the master visual attensons. The first study will evaluate the impact of including a master advisory attenson on crew detection, response, and error rates. If it is found that a master advisory attenson is required, it will be determined whether it has to be located within the pilot's 15° field of vision, or if it can be incorporated on the central information display, which is within the pilot's 30° field of vision.

The second study will evaluate the relative advantages and disadvantages of a combined master warning and caution attenson versus separate attensons. The combined attenson consists of a split legend switch-indicator that is labeled warning on top, and caution on the bottom. When a warning occurs, the top half of the switch-indicator illuminates red; when both a warning and a

Table 7.2.1-1. Visual Master Attenson

Variable	Concept A	Concept B
Number	Two	*
Location	Near 15-deg cones	*
Flash	No	*
Brightness	15 to 150 ft L	*
Size	1 deg	*
Cancellation logic	Manual and automatic	*
Duty cycle	N/A	*
Color	Red, yellow	*

Table 7.2.2-1. Aural Master Attenson

Variable	Concept A	Concept B
Number	Three	*
Signal-to-noise ratio	5 to 10 dB	*
Cancellation	Manual and automatic	*
Stereo type alerts	No	*
Duty cycle	N/A	*
Spectral character	In guidelines	*
Location	90 deg	*
Masking	Controlled via design	*

*Variable is identical for both concepts.

caution occur, both halves of the attenson illuminate. Warnings, cautions, or both, can be cancelled by the pilot by depressing this one switch-indicator. With separate attensons, two switch-indicators are used, one for warnings and the other for cautions. The only difference occurs when simultaneous warning and caution alerts are activated. With the combined master, only a single button depression is required to extinguish the alerts; with the separate attensons two button depressions are required. The results of these studies will be incorporated into both concepts for testing in Phase III.

7.2.2 MASTER AURAL ATTENSON(S)

Table 7.2.2-1 shows the variables associated with the master aural attenson for concepts A and B. Three aural masters (warning, caution and advisory) were included in the concepts to supplement the master visual attensons in alerting the crew and to provide information to direct crew responses. To be easily detectable, but not so loud as to startle the crew, the aurals were selected to be 5 to 10 db above the cockpit's ambient noise; the intensity of the aural alerts will be adjusted automatically as the ambient noise level of the cockpit changes. The aural alerts are cancelled in the same manner as the visual attensons (i.e., depressing the master visual attenson switch light cancels both the visual and aural signals); both are automatically cancelled when the problem has been resolved.

To avoid possible confusion of the aural attensons with the aural alerts existing in today's cockpits data obtained in Phase I will be used to select unique tones for each priority level. The warning tone will be intermittent or wavering, and will contain both high and low frequency components. The caution alert will be constant and midrange in frequency. The advisory tone will be a single-stroke, low frequency tone. To maximize perceived loudness and to enhance signal detectability, midfrequency tones (2000 to 4000 Hz) in combination with widely spaced frequency tones between 250 and 4000 Hz will be used.

To enhance detectability, loudspeakers for the aural alerts will be separated 90^{0} from other loudspeakers in the cockpit and the spectral characteristics of the aural alerts will be designed to minimize ambient noise masking effects.

A side study will be performed to evaluate the desirability of a master aural attenson for advisories; the potential gain in crew detection and response will be compared against the possible distracting effects of a third master tone. Again, the results of this study will be incorporated into the two alternative concepts.

7.2.3 VISUAL INFORMATION DISPLAY

The elements of the visual information display and the selected values/options for both concepts are shown in Table 7.2.3-1 A central display located on the pilot's main instrument panel will be used to present alphanumeric information that will specify the nature of the alert (e.g., YAW DAMPER FAIL). The display will be multicolor; warnings will be red, cautions yellow, and advisories blue. The display will be formatted by category and chronology. warnings will appear on the top of the display, cautions below warnings, advisories below cautions; the most recent alert will appear at the top of the alert category with the previous alerts waterfalling down.

The display will have the capability of presenting twelve alerts. If more than twelve alerts are active (i.e., not stored) the oldest and lowest priority alerts (e.g., advisory) will be removed from the display, and stored automatically on a second alert "page". An overflow indicator on or adjacent to the central display will illuminate, and the alerts on the second page will be manually selectable by depressing a page switch. Only two current alert pages will be provided. The pilot will have the capability to store any caution or advisory alert that he cannot resolve at the time. Warnings will not be removable from the display except when they have been resolved. A cursor control will be provided to store alerts selectively, and a recall button will be provided to re-display stored alerts.

To optimize readability, the brightness of the information display will be set between 15 and 150 ft L depending upon the ambient light level in the simulation facility. The alphanumeric characters will subtend 14 arc-minutes of visual angle; character spacing will be 7 arc-minutes.

Table 7.2.3-1. Information Display

Variable	Concept A	Concept B
Location	Central display	*
Format	Priority and chronological	*
Overflow	Paging	*
Store-recall	Yes, except for warnings	*
Brightness	15 to 150 ft L	*
Cues and aids	Box, arrow, etc.	*
Content	Short phrase (syntax)	*
Character size	14 min	*
Character spacing	7 min	*
Legibility	In literature	*

Table 7.2.4-1. Verbal Alerts

Variable	Concept A	Concept B
Type	Warnings	Warnings and caution elective
Format	Phrase	*
Model (M/F)	Female	*
Inflection	Monotone	*
Masking	Controlled by design	*
Repetition	Yes	*
Cancellation	Manual	Manual switch
Content	Status	*
Signal-to-noise ratio	5 to 10 dB	*
Multiple alerts	In sequence with repetition	No
Store-recall	No	Yes
Spectral character	Guidelines	*
Location	90 deg	*

*Variable is identical for both concepts.

Special symbols and characters will be used to aid the crew in detecting and responding to alerts. A flashing box around the alert message itself will be used to indicate newly displayed alerts, and a cue will be provided to indicate that more than twelve alerts are active and an overflow situation exists.

The content of the alphanumeric messages will be short phrases, with the location of the system presented first, the name of the system second, and the condition third (e.g., LEFT PACK FAIL OR RIGHT BRAKE HOT).

A side study will be performed to identify additional interactive functions (e.g., menu selection, checklist procedures) that should be added to the central display to facilitate crew performance. The results of this study will be incorporated into the system concepts for test and evaluation during later Phase III testing.

7.2.4 VOICE INFORMATION SYSTEM

Table 7.2.4-1 presents the variables and selected options for the voice information system for concepts A and B. The major differences between concepts A and B are:

- alert levels annunciated
- manner in which alerts are presented
- procedure for repeating alerts
- capability of storing/recalling alerts

For concept A, warnings only are annunciated verbally; the warnings are verbalized automatically, and repeat until the crew cancels them manually or until the problem is resolved. In concept B, both warnings and cautions can be verbally annunciated at pilot option. The attenson signals the onset of an alert; the pilot then decides whether to read the message on the information display, or to actuate a switch for the verbalized message.

The remaining differences between the two concepts are concept B's capability to recall the latest alert, and its procedure for handling multiple alerts

(verbal alerts cannot be stored or recalled in concept A). In concept B the pilot can recall the latest alert by depressing the switch. However, since only one alert is stored in concept B, no provision exists for recalling multiple alerts; multiple alerts are verbalized in sequence and repeat until cancelled. All of the other verbal alert variables for concepts A and B are identical.

Verbal messages will be presented in short phrases; message syntax will be identical to that used on the information display. A monotone female voice model will be used to present the messages, and the content of the messages will be advisory/status rather than command. The signal to noise ratio of the verbal message will be 5 to 10 db above ambient cockpit noise; spectral characteristics will be maximized to enhance message detectability. The loudspeakers for the verbal messages will be the same as those used for the master aural alerts, and will be located 90^{o} away from other loudspeakers in the cockpit.

Two voice information system side studies will be conducted. In the first study, the relative effectiveness of tone versus voice attensons will be evaluated. For this study the tones selected for warnings, cautions and advisories will be compared to the voice attensons "WARNING", "CAUTION", and "ADVISORY" to determine which results in faster detection and response times and lower error rates. The second study will determine whether the voice information system should verbalize all current alerts, or just say "MULTIPLE ALERTS" and direct the pilot to the central information display. Again, the results of these studies will be incorporated into both concepts for further testing in Phase III.

7.3 BASELINE ALERTING SYSTEM

To prepare for the validation effort in Phase III it was necessary to define a conventional alerting system which would be used as a benchmark to evaluate the candidate systems. This baseline system is not meant to represent any specific aircraft but rather was designed using the most common features existing in the U.S. built transport aircraft currently in service. The objective was to design a representative system to provide realistic data on conventional alerting concepts.

This approach resulted in a system which has a master visual attenson switch-indicator mounted in the glareshield. The top half will illuminate red with the legend "WARNING" for warning the bottom half lighting amber with the legend "CAUTION". The system will also have a centralized annunciator panel which will contain fixed legend lights to indicate the system which has a fault. These lights will be color coded in accordance to their level of urgency. Pressing the master visual attenson will cancel any of these lights that are illuminated. There will also be a cancel/recall function on the annunciator panel for those alerts which can be cancelled. The system will also have annunciator lights distributed over the front panel for such things as fire, auto-pilot disconnect, altitude alert, gear problems, configuration etc.

Dedicated aural tones will be used as they are in todays aircraft. The only voice messages that will be used will be from the Ground Proximity Warning System. The aural alerts that are cancellable will be cancelled by dedicated switches. A summary of the baseline system can be seen in Table 7.3-1.

7.4 TIME CRITICAL ALERTS

One area that must be treated in developing guidelines for an advance alerting system is that of extremely time critical warnings. The candidate alerting systems described in the previous section may not be able to provide quick enough response times for some situations which require immediate and accurate response. They are a very special set of operational or systems conditions requiring immediate attention. (e.g., ground prox, collision avoidance, windshear, etc.). They are defined as requiring unconditionally immediate corrective or compensatory action by the crew. The alerting system must therefore be able to accommodate this type of alert. However, implied in this definition is a high pilot confidence in the accuracy of the alert. To assure this confidence, it was necessary to assume that the system component reliability is high and the frequency of false alarms is very low. For this assumption to be met, the system designer will have to impose stringent requirements for those alerts that require the time-critical designation.

Table 7.3-1. Baseline Configuration

	Yes/no	Location	Color/tone	Flash
Master warning	Yes	Glareshield	Red/no	No
Single master				
Split legend				
Master caution	Yes	Glareshield	Amber	No
Central display				
Fixed legend	Yes	Central panel	Red, amber, blue	Yes—without master
Alphanumeric	Yes		Red, amber, blue	No—with master
Monchrome	No			
Distributed annunciators	Yes	N/A	Red, amber, blue	No

Additional features:

Dedicated tones:	Electronic/mechanical mix; also include SELCAL, CREW CALL; cancellable, dedicated cancel switches
Voice message:	GROUND PROX, cockpit speaker environment
Inhibits:	No
Cancel/recall:	Yes; can cancel masters and central display, cannot cancel distributed annunciators

It is anticipated that with advancements in sensor technology the possible number of time-critical alerts will increase. With the addition of appropriate sensors it may be possible to include such things as collision avoidance, windshear alerts, and flight guidance after a major failure or catastrophic flight situation. With this kind of alert evolving, it became apparent that a systems approach to crew alerting must have provisions to accommodate them.

The questions that were posed concerning time critical alerts were as follows:

1. Does the time-critical alert require a separate display? If so, where should it be located?

2. Is graphic presentation of the data more efficient than alphanumeric presentation? Is a combination of graphic and alphanumeric data an appropriate presentation format?

3. Should voice be used with the visual presentation?

4. Should the alert provide only situational information or should it provide guidance in making the appropriate response?

These questions will be treated in the design and conduct of the Phase III testing.

7.5 FLIGHT ENGINEER'S ALERTING SYSTEM

Another area that will be treated in the Phase III tests will be the validation of the alerting system concept with respect to the Flight Engineer. When the the crew alerting system must also accommodate the additional crewmember's functions.

The flight engineer will be provided with an information display comparable to the one on the main instrument panel and will permit alerting information to be located in a single location. The display will be programmable and repeat the messages that appear on the pilot's display; it will also have controls to

enable the flight engineer to interact with the system. These controls will include: line select keys so that any line or message on the display can be designated; store and recall switches permitting the display to be cleared with the ability to recall the alert messages; a checklist switch which, when used in conjunction with line selection will allow procedural/checklist information to be presented for the selected fault; a page key and rocker switch which permit the user to review more than one page of procedural information or messages (see figure 7.5-1). This presentation concept will be compared with a conventional flight engineer's station in Phase III of the program.

7.6 FOLLOW ON VERIFICATION AND EVALUATIONS

Phase III of the program will have as its objectives:

The resolution of the questions generated during the definition of the candidate systems;

Validation of the candidate systems, including the flight engineers station;

Identification of the presentation media for time-critical alerts;

Development of system design recommendations and guidelines;

Development of testing procedures for designer to assure that his system will be compatible with the flight deck environment for all phases of flights.

Phase III will implement the candidate concepts in simulator hardware, validate them, and evaluate the effects of time-critical alerts and the flight engineer's station. The part task tests identified in Phase I will be conducted in conjunction with construction of the simulator hardware so that this data can be included in the final system simulation.

System validation will treat the two pilot's candidate system concepts and the second officer candidate system. These evaluations will compare the advanced alerting concepts with a representative conventional system (see figure

7.6-1). The advanced concepts will be validated as viable alternatives if crew performance with these systems is equal to or better than their performance on the baseline system.

The tests using the time-critical alerts will be designed to answer the questions identified in section 7.4.

REFERENCES

Adams, J. A., and Chambers, R. W., Response to Simultaneous Stimulation of Two Sense Modalities, Journal of Experimental Psychology, Volume 63 PP. 193-206, 1962.

Adams, J. A., Humes J. M., and Stenson, H. H., Monitoring of Complex Visual Displays: III Effects of Repeated Sessions on Human Vigilance, Human Factors, Volume 4(3), PP. 149-158, 1962.

ARP 450D, Flight Deck Warning, Caution and Advisory Signal System, Society of Automotive Engineers, Inc., Draft Copy, September, 28, 1979.

Bate, A. J. Cockpit Warning Systems Comparative Study Report No. AMRL-TR-68-193, Aeromedical Research Laboratory, Wright-Patterson AFB, Ohio, 1969.

Boucek, G. P., Veitengruber, J. E., and Smith, W. D., Aircraft Alerting Systems Criteria Study, Volume II: Human Factors Guidelines and Aircraft Alerting Systems. FAA Report, FAA-RD-76-222, May 1977.

Buckner, D. N., and McGrath, J. J., A Comparison of Performance on Single and Dual Sensory Mode Vigilance Tests, Technical Report, NR 153-199, (Office of Naval Research) Human Factors Research, Inc., Los Angeles, California, 1961.

Cherry, E. C., Source Experiments on the Recognition of Speech with One and with Two Ears, Journal of the Acoustical Society of America, Volume 25, PP. 975-979, 1953.

Christ, R. E., and Teichner, W. H., Color Research for Visual Displays, Final Report Janair 730 703, July, 1973.

Christ, R. E., Four Years of Color Research for Visual Displays, Proceedings of the Human Factors Society, 21st Annual Meeting, San Francisco, California, October 17-20, 1977.

Christ, R. E., and Corso, G. M., Color Research for Visual Displays, Final Report, ONR-CR 213-102-3, July 1975.

Christ, R. E., Stevens, A. C., and Stevens, D., Color Research for Visual Displays, Final Report Janair 741103, July 1974.

Coates, G. D., Loeb, M., and Allussi, E. A., Influence of Observing Strategies and Stimulus Variables on Watch Keeping Performances, Ergonomics, Volume 15, PP. 379-386, 1972.

Cooper, G. E., A Survey of the Status of and Philosophies Relating to Cockpit Warning Systems, Report No. NASA CR-152071, NASA Ames Research Center, Moffett Field, California, 1977.

Crawford, A., The Perception of Light Signals: The Effect of Mixing Flashing and Steady Irrelevant Lights, Ergonomics, Volume 6, PP. 287-294, 1963.

Crawford, A., The Perception of Light Signals: The Effect of the Number of Irrelevant Lights, Ergonomics, Volume 5, PP. 417-428, 1962.

Davenport, W. G., Stimulus Variables in Vigilance: Signal Duration and Intensity, Austrailan Journal of Psychology, Volume 20(2), PP. 129-133, 1968.

Davis, R. C., Motor components of Responses to Auditory Stimuli: The Relation of Stimulus Intensity and Instructions to Respond, American Psychologist, Volume 2, PP. 308, 1947.

Edwards, R. A., The Perception and Application of Flashing Lights, Proceedings of the Interrnational Symposium, Imperial College of Science and Technology, London, England, PP. 209-216, April 19-22, 1971.

Egan, J. P., Carterette, E. C., and Thwing, E. J., Some Factors Affecting Multi-Channel Listining, Journal of the Acoustical Society of America, Volume 26, PP. 774-782, 1954.

Erickson, J. B., Voice Warning Questionnaire Results, Internal McDonnell-Douglas AVI, Dated 12/12/78.

Fletcher, H., and Munson, W. A., Loudness, its Definition, Measurement, and Calculation, Journal of the Acoustical Society of America, Volume 5, PP. 82-108, 1933.

Gerathewohl, S. J., Conspicuity of Steady and Flashing Light Signals: Variation of Contrast, Journal of the Optical Society of America, Volume 43, PP. 567-571, 1953.

Haines, R. F., Dawson, L. M., Galuan, T., and Reid L. M., Response Time to Colored Stimuli in the Full Visual Field, Report No. NASA TN D-7927, National Aeronautics and Space Administration, Washington, D.C., 1975.

Haines, R. F., and Gililand, K., Response Time in the Full Visual Field, Journal of Applied Psychology, Volume 58(1), PP. 289-295, 1973.

Haines, R. F., Gross, M. M., Nylen, D., and Dawson, L. M., Peripheral Visual Response Time to Colored Stimuli Imaged on the Horizontal Meridian, Report No. NASA TMX-30 86, NASA Ames Research Center, Moffett Field, California, 1974.

Hart, S. G., and Simpson, C. A., Effects on Linguistic Redundancy on Synthesized Cockpit Warning Message Comprehension and Concurrent Time Estimation, Proceedings of the Twelfth Annual Conference on Manual Control, NASA TMX-73, PP. 309-321, 1976.

Hill, N. E. G., The Recognition of Colored Light Signals which are near the Limit of Visibility, Proceedings of Psychological Society (London), Volume 59. PP. 560, 1947.

Howarth, E. I., and Ellis, K., The Relative Intelligibility Threshold for One's Own Name Compared with other Names, Quarterly Journal of Experimental Psychology, Volume 13, PP. 236-239, 1961.

Kemmerling, P., Geiselhart, R., Thorburn, D. E., and Crowburg, J. G., A Comparison of Voice and Tone Warning Systems as a Function of Task Loading, Technical Report ASD-TR-69-104, Air Force Systems Command, Wright-Patterson AFB, Ohio, 1969.

Kerle, E. W., Intelligibility Testing of Voice Model and Phoneme-Synthesized Voices for Aircraft Caution-Warning Systems, California State University, Long Beach, California, 1979.

Klemmer, E. T., Time Sharing Between Frequency-Coded Auditory and Visual Channels, Journal of Experimental Psychology, 55, pp. 229-235, 1958.

Kohfeld, D. L., Simple Reaction Time as a Function of Stimulus Intensity in Decibels of Light and Sound, Journal of Experimental Psychology, Volume 88(2), PP. 251-257, 1971.

Kuess, P. J. G., Reaction Time to the Second of Two Sortly Spaced Auditory Signals Both Varying in Intensity, ACTA Psychological, Volume 36, PP. 226-238, 1972.

Luckiesk, M., Light, Vision and Seeing. Van Nostrand, New York, 1944.

McCormick, E. J., Human Factors Engineering, McGraw-Hill Book Company, New York, 1970.

Meister, D., and Sullivan, D. J., Guide to Human Engineering Design for Visual Displays, AD 693237, Office of Naval Research, Department of the Navy, Arlington, Virginia, 1969.

Miller, G. A., The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information, Psychological Review, Volume 63(2), PP. 81-96, 1956.

Mills, A. W., On the Minimum Audible Angle, Journal of the Acoustical Society of America, Volume 30, PP. 237-246, 1958.

MIL-SPEC-18012B, Markings for Aircrew Station Displays, Design and Configuration of, Department of Defense, February 17, 1972.

MIL-STD-411D, Aircrew Station Signals, Department of Defense, August 30, 1974.

MIL-STD-1472B, Human Engineering Design Criteria for Military Systems, Equipment and Facilities, Department of Defense, May 10, 1978.

Morgan, C. T., Cook, J. S., Chapanis, A., and Lund, M. W., (eds), Human Engineering Guide to Equipment Design, McGraw-Hill Book Company. New York, 1963.

Pollack, I., The Information of Elementary Auditory Displays, Journal of the Acoustical Society of America, Volume 24, PP. 745-749, 1952.

Pollack I, and Ficks, L., Information of Multidimensional Auditory Displays. Journal of the Acoustical Society of America, Volume 26, PP. 155-158, 1954.

Pollack, J. D., Reaction Time to Different Wavelengths at Various Luminanies, Perception and Psychophysics, Volume 3, PP. 17-24, 1968.

Pollack, J. D., and Tecce, J., Speech Annunciator Warning Indicator System: Preliminary Evalucation, Journal of the Acoustical Society of America, Volume 30, PP. 58-61, 1958.

Raab, D., and Fehrer, E., The Effects of Stimulus Duration and Luminance on Visual Reaction Time, Journal of Experimental Psychology, Volume 64(3), PP. 326-327, 1962.

Randle, R. J., Larsen, W. E., and Williams, D. H., Human Factors Guide to Evaluating Cockpit Alerting and Warning Systems, NASA Ames Research Center, Moffett Field, California, 1975.

Reynolds, R. E., White, R. M., Jr., and Hilgendorf, R. L., Detection and Recognition of Colored Signal Lights, Human Factors, Volume 14, PP. 227-236, 1972.

Senders, V. L., The Effect of Number of Dials on Qulaitative Reading of a Multiple Dial Panel, WAOC Technical Report 52-182 Wright Air Development Center, United States Air Force, Wright-Patterson AFB, Ohio, 1952.

Sheehan, D. J., Heads-Up Display Warning Requirements Research, Final Report NR 213-086, Office of Naval Research, Department of the Navy, Arlington, Virginia, 1972.

Shower, E. G., and Biddolph, R., Differential Pitch Sensitivity of the Ear, Journal of the Acoustical Society of America, Volume 3, PP. 275-287, 1931.

Siegel, A. I., and Crain, K., Experimental Investigations of Cautionary Signal Presentations, Ergonomics, Volume 3, PP. 339-356, 1960.

Simpson, C. A., Effects of Linguistic Redundancy on Pilot's Comprehension of Synthesized Speech. Proceedings of the Twelfth Annual Conference on Manual Control, NASA TMX-73, PP. 294-308, May, 1976.

Simpson, C. A., and Williams, D. H., Human Factors Research Problems in Electronic Voice Warning System Design, N75-33681, 11tn Annual Conference on Manual Control, NASA Ames Research Center, Moffet Field, Calif., May, 1975.

Simpson, C. A., and Hart, S. G., Required Attention for Synthesized Perception for Three Levels of Linguistic Redundancy, 93rd Meeting of the Acoustical Society of America, Pennsylvania State College, June, 1977.

Simpson, C. A., and Williams, D. H., The Effects of an Alerting Tone and of Semantic Context on Pilot Response Time for Synthesized Speech Voice Warnings in a Simulated Air Transport Cockpit, MCI Report No. 78-001, NASA Ames Research Center, Moffett Field, California, 1978.

Society of Automotive Engineers, Aerospace Recommended Practice: Flight Deck Visual, Audible and Tactical Signals (ARP-450C), Society of Automotive Engineers, Inc., New York, December, 1970.

Society of Automotive Engineers, Aerospace Recommended Practice: Flight Deck Visual, Audible and Tactical Signals (Draft ARP-450D), Society of Automotive Engineers, Inc., New York, September, 1979.

Speith, W., Curtis, J. F., and Webster, J. C., Responding to One of Two Simultaneous Messages, Journal of the Acoustical Society of America, Volume 26, PP. 391-396, 1954.

Steinman, A. R., Reaction Time to Change Compared with other Psychophysical methods, Archives of Psychology, New York, Volume 292, PP. 34-60, 1944.

Steinman, A. R., and Venias, S., Simple Reaction Time to Change as a Substitute for the Disjunctive Reaction, Journal of Experimental Psychology, Volume 34, PP. 152-158, 1944.

Stevens, S. S., Handbook of Experimental Psychology, John Wiley and Sons, Inc., New York, 1951.

Vancott, H. P., and Kinkade, R. G., Human Engineering Guide to Equipment Design, United States Printing Office, Washington, D.C., 1972.

Vanderschraaf, A., Problem Area: Warning Systems, Fokker, VFW Aircraft, Proceedings from the 20th International Air Safety Seminar of the Flight Safety Foundation, Anaheim, California, October 25-29, 1976.

Veitengruber, J. E., Boucek, G. P., and Smith W. D., Aircraft Alerting Systems Criteria Study, Volume I: Collation and Analysis of Aircraft Alerting Systems Data. FAA Report, FAA-RD-76-222, May, 1977.

Weingarten, F. S., Wavelength Effect on Visual Latency, Science, Volume 176, PP. 692-694, 1972.

White, W. J., and Schneyer, S., Pocket Data for Human Factors Engineering, Report No. VB-1227-V-3, Cornell Aeronautical Laboratory, Inc., Buffalo, New York, 1960.

Tannas, L. E., Jr., and Goede, W. F., Flat Panel Displays, a Critique, I.E.E.E. Spectrum, July, 1978. PP 26-32.

APPENDIX A

TEST FACILITIES DETAILED DESCRIPTION

APPENDIX A

A.1 THE BOEING COMPANY TEST SIMULATOR LABORATORY

The laboratory includes special sound, lighting, light control measure, projection screens, a control room and electronics, and is located in the Renton (Washington) Simulation Center, Bldg. 10-65. A layout is shown in Figure A.1-1.

Lighting - Both red and white area lighting is installed in the lab, and is capable of providing discrete area lighting control of levels up to 100 fl; for brighter test conditions, banks of manually-positioned, high-brightness lights providing up to 10,000 fl are also available.

Test in Progress Signs - are positioned at all entry doors to prevent test interruption.

Light Exclusion and Control - All access doors, windows and ceilings are finished and weatherstripped to prevent light leakage.

Painting - The entire room is painted with DePolo blue which gives a pleasing appearance when illuminated but turns black at reduced light levels.

Electric Power - 115 volt, 15 amp, 60 Hz power is provided to 15 locations in the area; 115 volt, 15 amp, 400 Hz, 3 phase power is also available.

Projection Screens are flat There are three 17' x 20' projection screens located 17 feet from the test subject's eyes and perpendicular to his line of sight.

Air Conditioning - Special airconditioning is used to cool the large quantities of electronics and projectors used in this area.

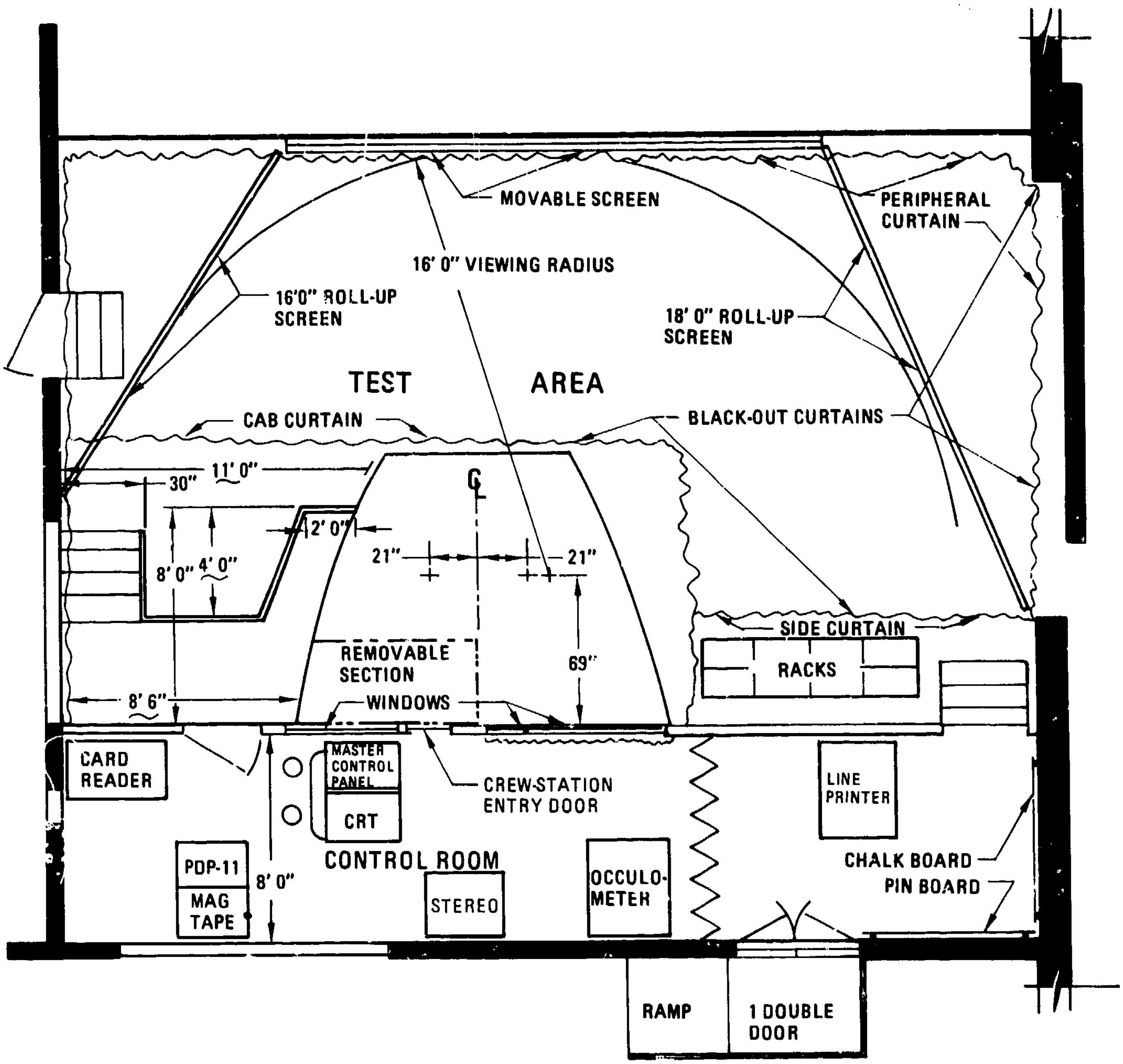


Figure A.1-1. Crew Systems Flight Deck Lab

Control Room - The control room houses the majority of the supporting electronic systems and has been configured to best allow the test conductor to monitor and control the test with direct vision or by closed circuit TV.

The key elements of the facility will be discussed in the following sections.

A.1.1 DEVELOPMENTAL "D" CAB

The Developmental Cab is unique in that it has been designed to accommodate either new airplane development or crew systems flight deck research. For the latter purpose. the cab is fitted with standard aircraft instruments, a number of alphanumeric displays, a side arm controller (on the right) and wheel/column control (on the left), a throttle mechanism, clock display, two keyboards communication headsets and cockpit speakers/microphones. These units are wired to the Switch Monitor and Light Driver (SMOLD), Digital to Analog and Synchro Converter, PDP 11 computer, or directly to the Master Control Panel (Figure A.1.1-1.).

A.1.2 MASTER CONTROL CONSOLE

This console contains the master control panel by which the experimenter monitors and controls the test and interfaces with the following systems: stereo, cab instrumentation computer (PDP11) and the projection system.

Contained within this unit are two F8 microprocessors, one dedicated to the cab's alphanumeric displays, the second controlling the stereo system. Also contained are interfacing circuits between the stereo, cab. data bus, PDP11 and the SMOLD (Figures A.1.2-1, -2, 3).

A.1.3 PDP11/10 MINI COMPUTER

The PDP11 computer is used to control the test and to provide all of the logic to drive the instrumentation; stored within its memory is the test scenario. Its main feature is a 16-bit parallel central processor unit (CPU) which performs all arithmetic and logical operations required in the system. Peripherals include a magnetic tape memory, high speed line printer, card

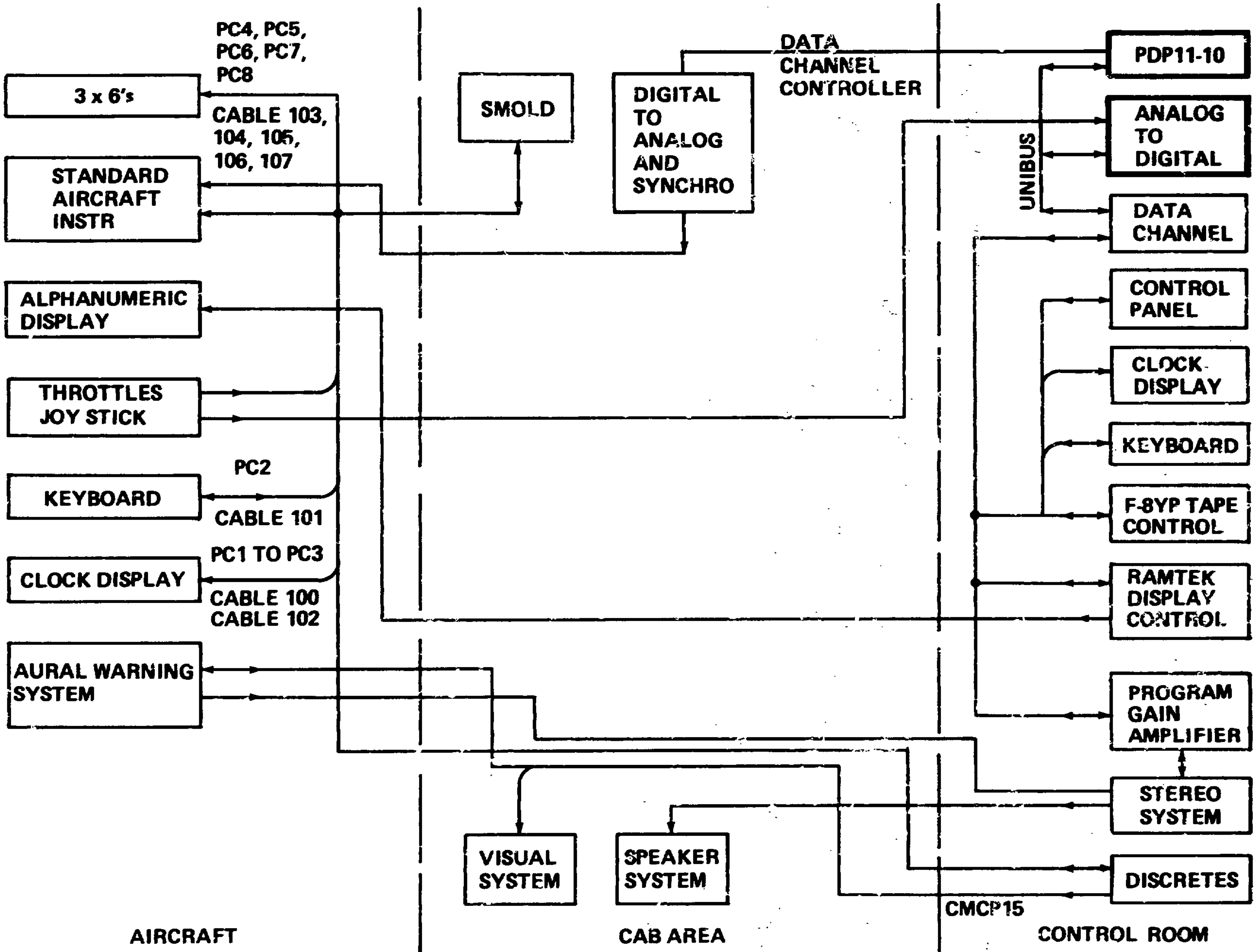


Figure A.1.1-1. Developmental Cab Interface Schematic

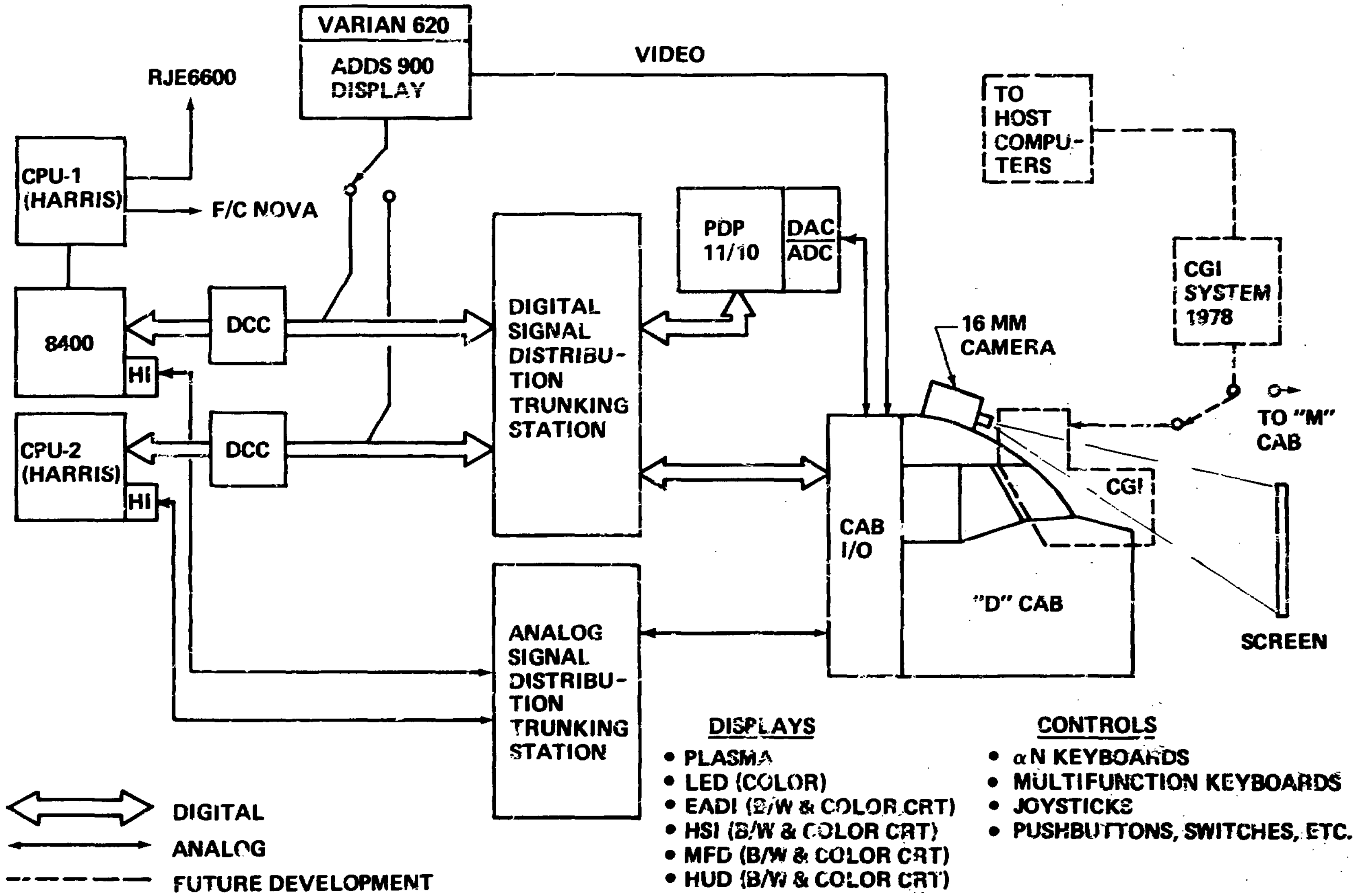


Figure A.1.2-1. "D" Cab Interface Diagram

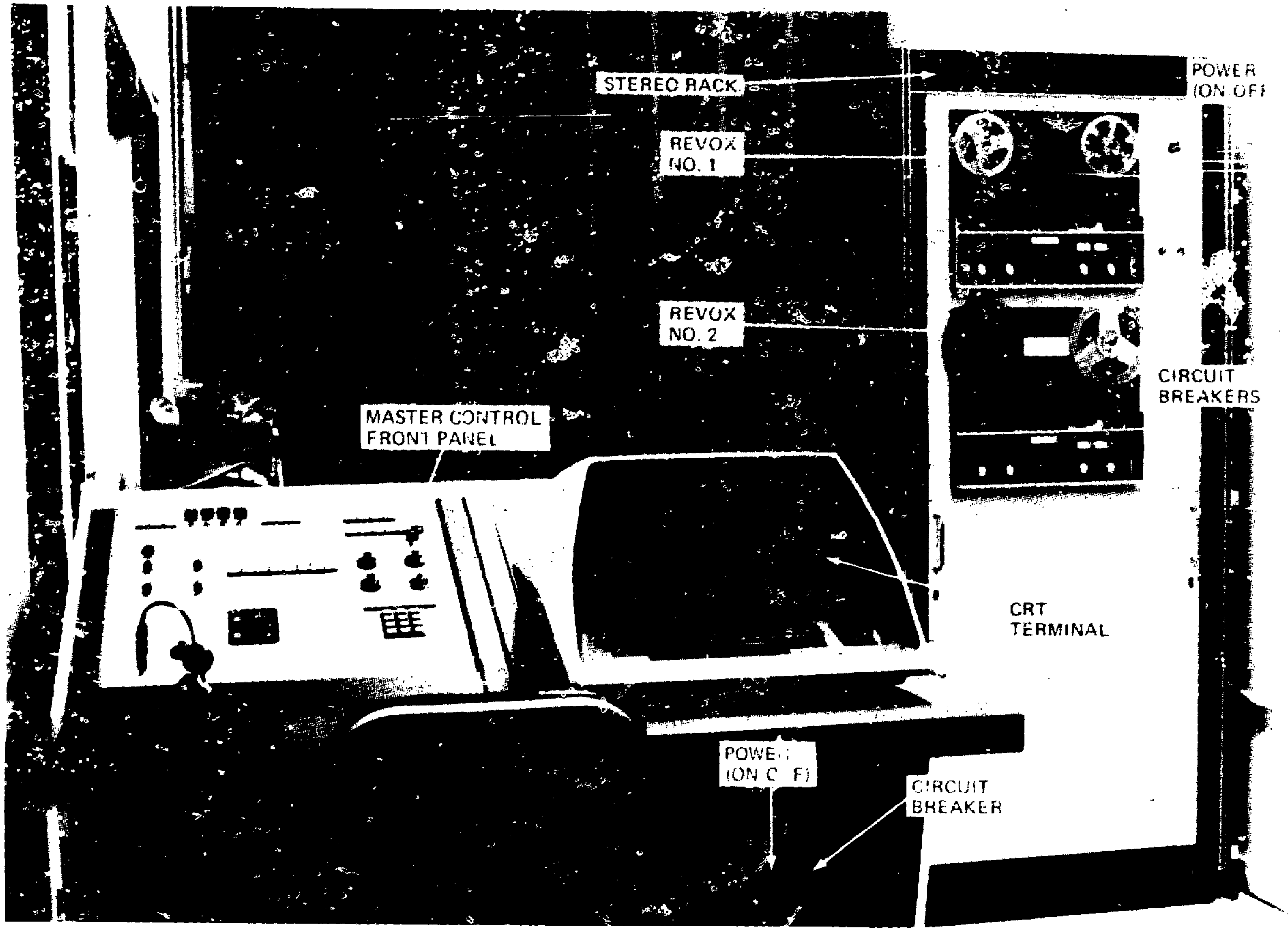


Figure A.1.2-2. Master Control Console (Front View)

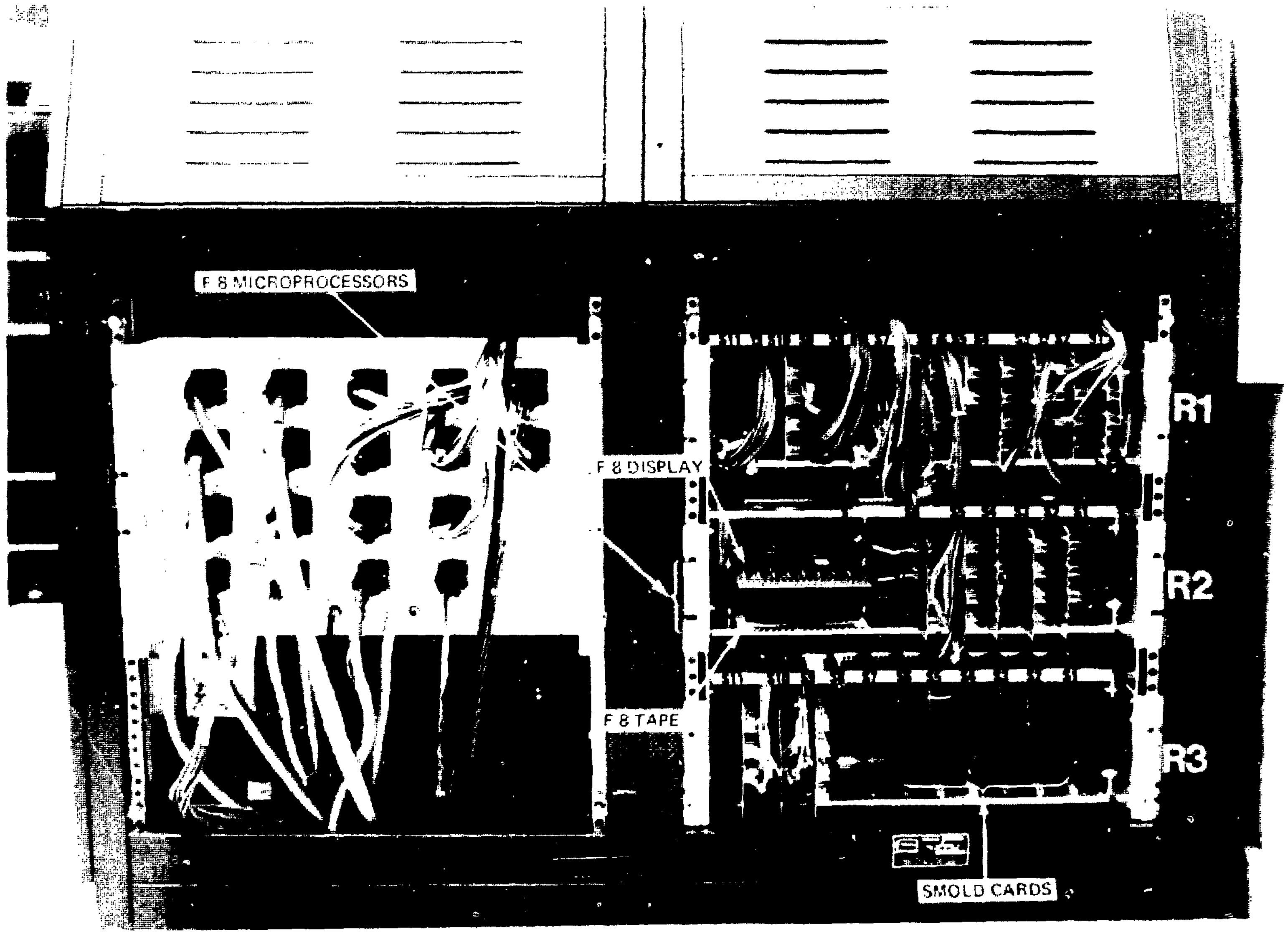


Figure A.1.2-3. Master Control Console (Rear View)

reader, high speed paper tape reader and punch, and a terminal alphanumeric display and keyboard.

A.1.4 VISUAL SYSTEM

A visual system comprised of three independent projection systems provides out-the-window visual tasks for the test subject. The left and right systems consist of two 35 mm Kodak slide projectors. The center projection system can be supported by either a 16 mm time and motion projector or a third slide projector. All systems can be controlled by the test conductor or by the PDP11 computer.

A.1.5 MAGNETIC TAPE RECORDING SYSTEM

A seven track magnetic tape recorder is used to record test milestones, time references, time lines of simulator state, crew performance, physiological data and other related equipment operations; provisions also exist to support oculometer instrumentation.

A.1.6 AUDIO/STEREO SOUND SYSTEM

This unit housing three independent sound systems independently recording First Officer/test conductor conversations, can be controlled from the Master Control Console or by the PDP 11.

Sound system number one has aero noise and engine noise recorded on separate tracks and is comprised of the following components: Revox stereo tape deck, equalizer, 300 watt amplifier and four large speakers.

Sound system number 2 has its own Revox stereo tape deck with separate recordings of voice/aural warnings and ATC/weather information; it utilizes a 120 watt amplifier to drive six speakers located at different places within the cab.

Communications within the lab between the test conductor, the oculometer operator, the test subject, and the speaker in the control room area are

driven by the Marantz amplifier, channeled through the Teac to provide the appropriate gain, and recorded on a Sony TC 138 cassette voice recorder (Figure A.1.6-1).

A.1.7 TV TEST MONITORING SYSTEM

The TV monitoring system used to monitor the test subject(s) in the cab, consists of four CONRAC vidicon cameras, camera controls and a 1" video tape recorder. One camera is used for overall monitoring, two are for close up monitoring of key displays and controls, and the fourth is trained on a digital clock; two of the four cameras are fixed with zoom capability. A 23" monitor and special effects generator to allow various combinations of split screen monitoring and recording control are located in the control room.

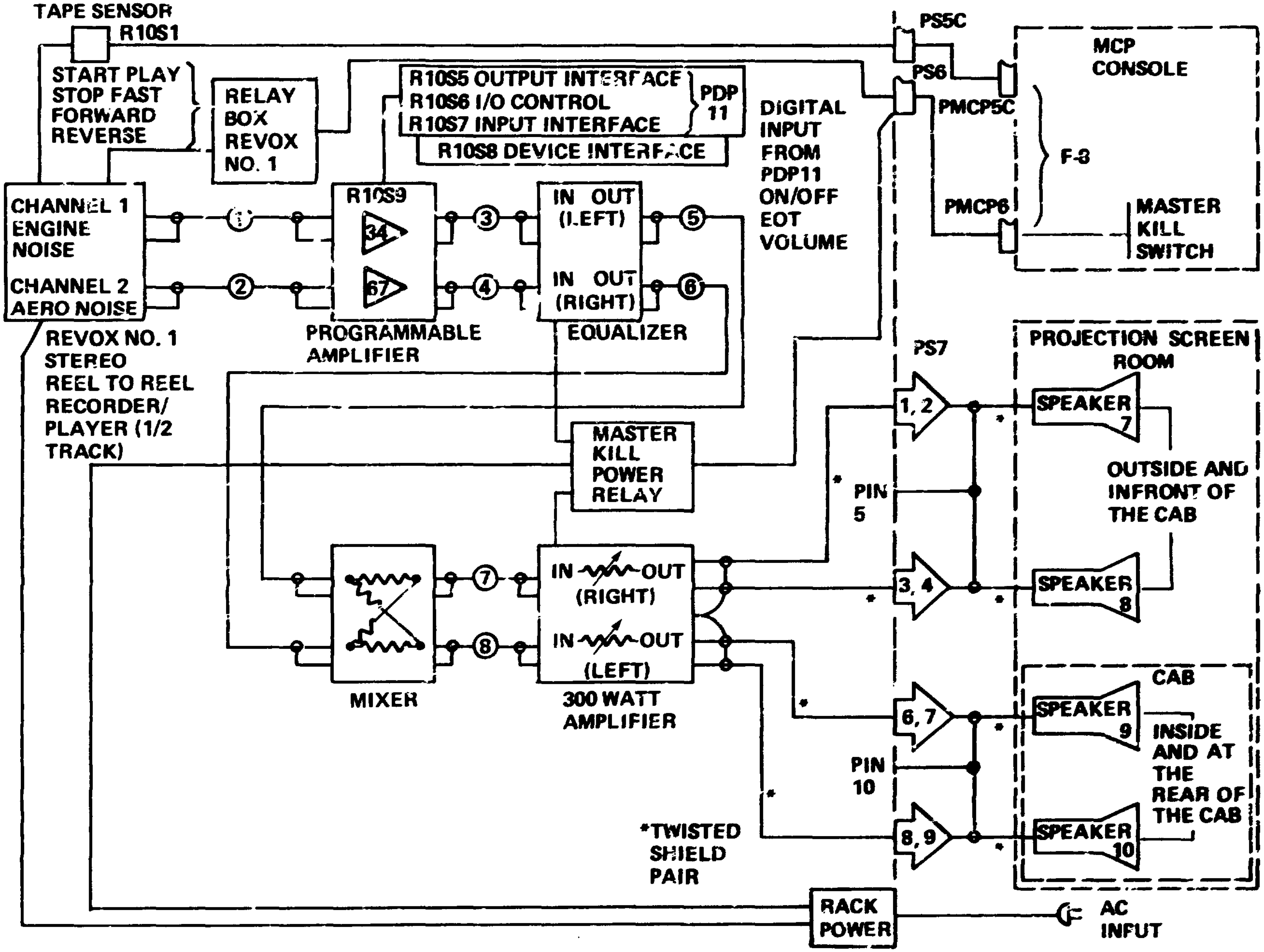


Figure A.1.6-1. Developmental Cab Audio System

A.2 DIGITAL EQUIPMENT TECHNOLOGY ANALYSIS CENTER (DETAC)

A.2.1 GENERAL DESCRIPTION

DETAC is a technology investigation facility used for the purpose of conducting studies and providing hands-on experience for engineers assigned to tasks associated with digital equipment. This facility was established to fulfill a requirement to upgrade the existing electronic system evaluation capabilities, particularly in the area of aircraft digital systems, inclusive of flight control computers and advanced display concepts. The facility as well as the cockpit fixture are illustrated in Figure A.2.1-1.

Aircraft that utilize several different types of digital computer systems require careful study of software structure, allocation of hardware/software function, redundancy management, etc., to obtain proper reliability and safety of flight. The DETAC system permits engineers to investigate problem areas, conduct real-time simulations, and monitor the design and integrity of vendor production hardware.

The DETAC facility has the following operational features:

- Central digital computer with a real-time operating system, 96K words of memory with memory mapping floating-point hardware, and cache memory.
- FORTRAN IV software package.
- Interactive graphics unit with FORTRAN level software package.
- Three satellite digital computers with 32K words of memory, 512-word writable control store, and floating-point firmware package.
- High speed, party-line communications link, 500K words/sec.
- Cockpit simulation apparatus and displays.

Figure A.2.1-1. Digital Equipment Technology and Analysis Center (DETAC)

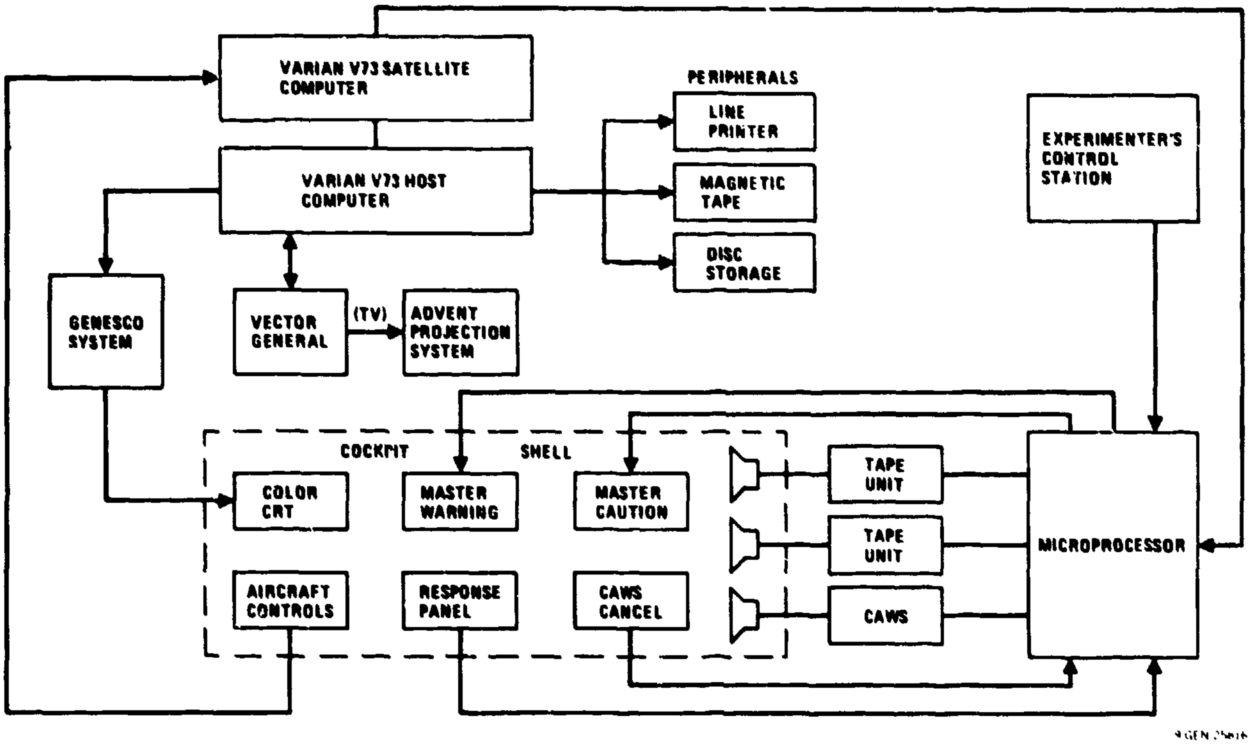


Figure A.2.2-1. DETAC Hardware Configuration for Tests III and IV of Aircraft Alerting Systems Standardization Study

The DETAC facility is used to support advanced commercial and military studies in digital flight controls, integrated cockpit technology, aircraft mutiplex systems, and advanced military tactical displays. Specific types of digital avionics investigations include system architecture and stability studies, digital autopilot evaluation and mechanization studies; higher-order language applications; hardware, software, firmware tradeoffs; display format studies; and software reliability and certifiability studies.

DETAC has five major elements: Central computer and peripherals, satellite computer and peripherals, interactive graphics, general input/output (I/O) hardware, and cockpit fixtures. There basic functional elements and associated equipment are shown in Figure

The Sperry Univac V76 minicomputer is general purpose and micro-programmable. A cache enhances memory access for faster operation. Three Sperry Univac V76 minicomputers provide computations and simulation support for the central computer. Peripheral support equipment includes a Dec-writer 111 terminal, card reader, magnetic tape, cassette tape, Century Data CDS-114 disc, Infotron Vistar/GT alphanumeric display terminal, and Varian Statos-31 printer/plotter.

Interactive computer programs define wind shears, turbulence, ILS characteristics, aircraft initial conditions, and flight-control-system parameters. Simulations of advanced flight-guidance system utilize the unique capability of DETAC to operate several computers asynchronously in parallel. Multiple computers simulate redundant avionics systems, while other computers simulate the head-up visual scene cockpit displays. Data-reduction programs plot selected simulation parameters on a Vector General display and the Statos electrostatic plotter.

This fixture consists of a wooden mockup of a three man wide body cockpit with crew seats, control mechanisms, CRT displays, instrument panel and supporting structure interfaced to the Varian computers. The DETAC cockpit can be easily reconfigured to permit the study of advanced cockpit concepts. The instrument panel consists of several color and black-and-white CRT's. A microprocessor provides a flexible interface between the cockpit controls and the satellite computers. Head-up and external visual displays are produced by an Advent television projection system.

A.2.2 FACILITY MODIFICATIONS FOR AIRCRAFT ALERTING SYSTEM STANDARDIZATION STUDY

Several facility modifications were made prior to Phase I testing. Figure A.2.2-1 represents a schematic diagram of the hardware configuration that was used. An experimenters control station was added as well as a microprocessor interface unit for activation of warning system devices in the cockpit. A prototype synthetic voice warning system was installed along with audio equipment for ambient noise simulation and ATC communications. In addition, a GENESCO color graphics system was added to drive the central caution and warning display unit. Figure A.2.2-2 illustrates the specific modifications made in the DEVAC cockpit. As can be seen, a modular overhead panel was installed to accommodate lighted switches to be used for simulated fault identification and correction. Three sets of speakers were also installed as part of the modification process. The overhead speakers were provided for the transmission of simulated ATC communications. The side panel speakers were used to introduce alert messages and the floor mounted speakers were employed to simulate ambient cockpit noise. A number of instrument panel configuration changes were also made. These modifications are shown in Figure A.2.2-3. Master warning and caution lights were added to the glareshield as well as a Central Aural Warning System cancel switch which functions to cancel and reset any on-going auditory alerts. A control display unit with its associated control keys is located at the center of the instrument panel while a hard wired annunciator matrix was installed in the Captain's primary field of view for required warnings and system status information in an operational system, the annunciator matrix would serve as a back up device to be used during control display unit failure. It is presented here for demonstration purposes and was not a part of the experimental tests. Figure A.2.2-4 shows the pilot's position relative to the two side panel speakers. As can be seen, his head is positioned between these two speakers while the ATC speaker is positioned directly above his head. The pilot's eye position corresponds roughly to the design eye reference point. Control of the aircraft is exercised by means of a side stick controller as illustrated in Figure A.2.2-5. The overhead response panel used to acknowledge fault messages as well as the glareshield mounted microphone used to record the pilot's verbal responses are also visible in Figure 4.2.2-5. The external visual scene was

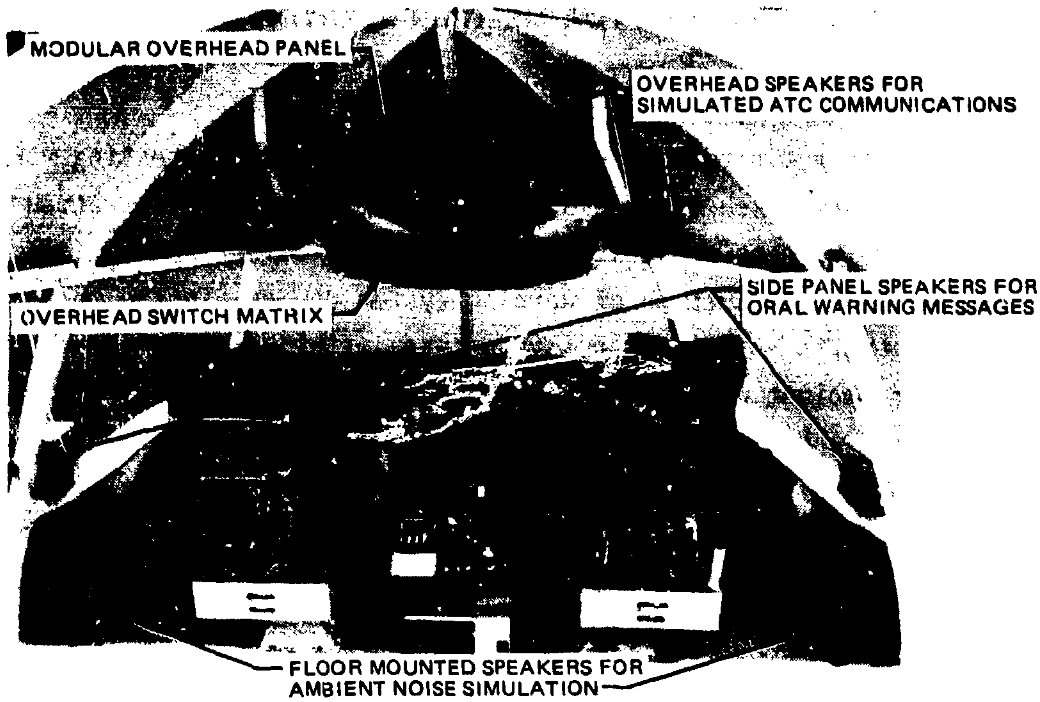


Figure A.2.2-2. Modification Made in DETAC Cockpit for Tests III and IV

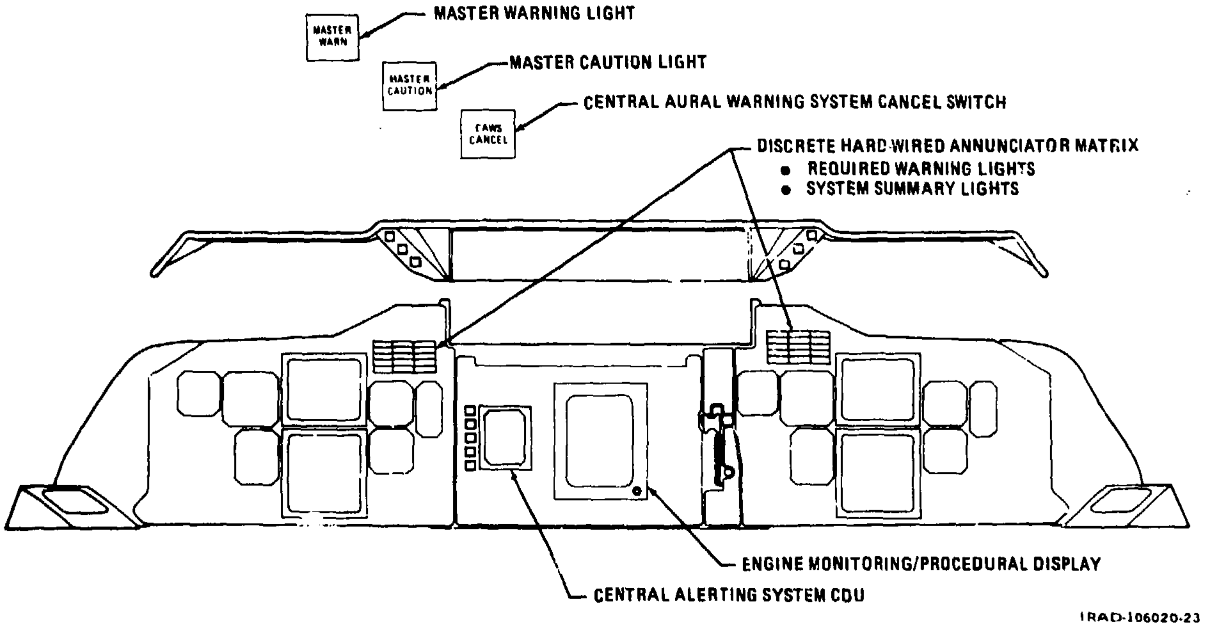


Figure A.2.2-3. Instrument Panel Configuration

Figure A.2.2-4. Pilot Position Relative to Side-Mounted Speakers

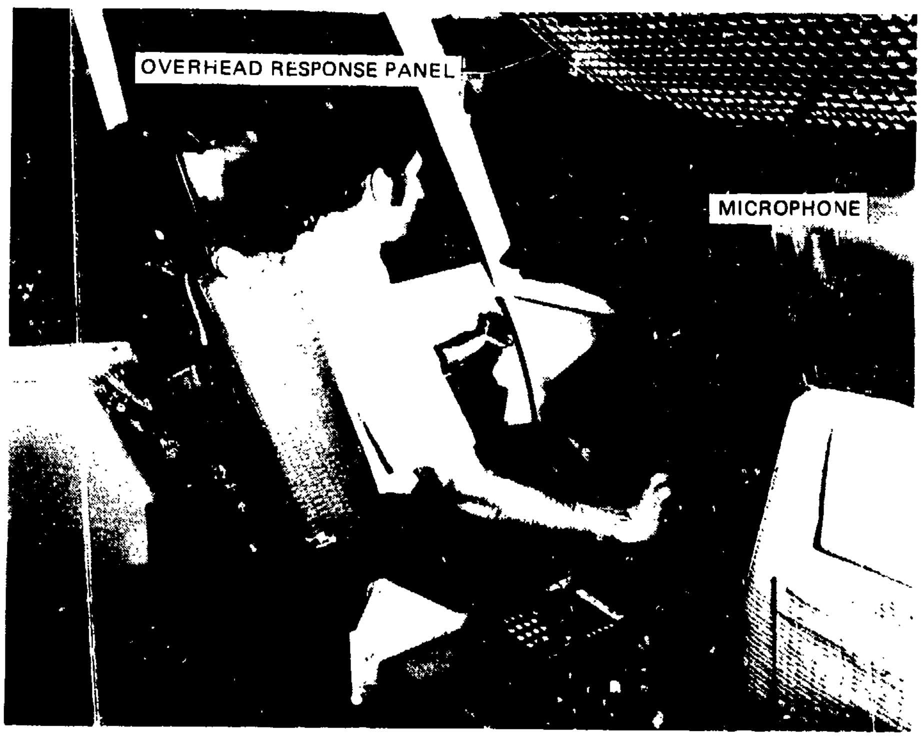


Figure A.2.2-5. Sidestick Controller, Overhead Response Panel, and Glareshield-Mounted Microphone

generated on a computer graphics terminal and presented directly in front of the pilot on an ADVENT projection screen at a distance of approximately 15 feet. The projection screen can be seen in relation to the cockpit fixture in Figure A.2.2-6. Head-Up Display (HUD) symbology was used by the pilot for visual guidance during simulated approaches. The pilot's task was to maintain the aircraft symbol centered over the command symbol. The difficulty of this two axis tracking task could be modified by introducing various levels of turbulence. A representation of the HUD symbology as used for Tests III and IV can be seen in Figure A.2.2-7.

As can be seen in Figure A.2.2-1, control of the experiments was maintained from a remote location within the DETAC facility. With this configuration it was possible to initiate each test trial and introduce the appropriate alert messages without having to enter the cockpit fixture. Video taping equipment was installed to record the pilot's movements and verbal responses for subsequent analysis.

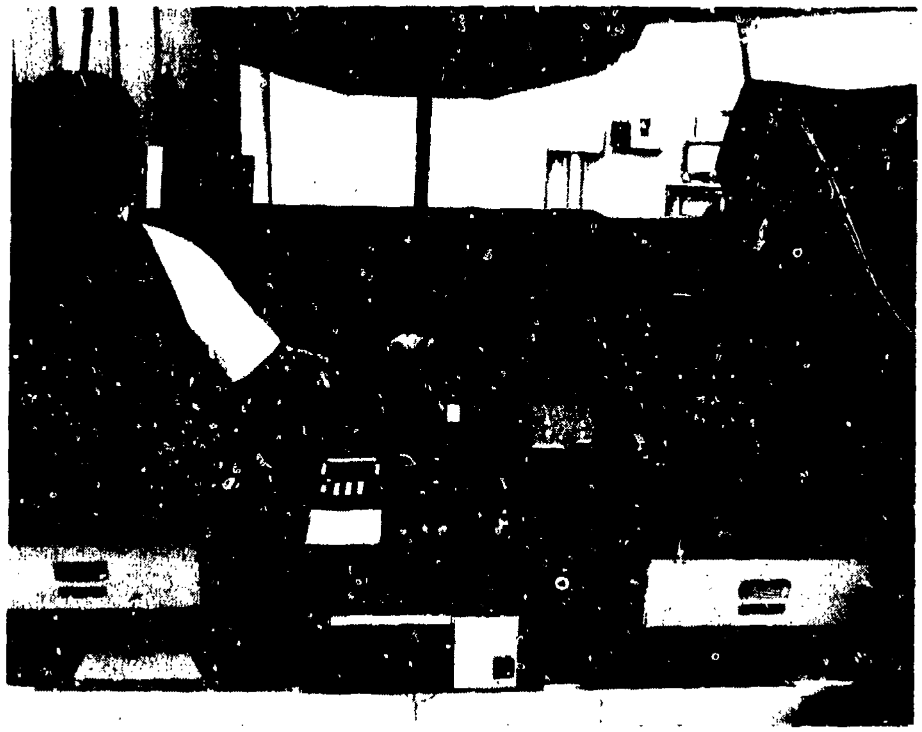


Figure A.2.2-6. Advent Projection Screen in Relation to Cockpit Fixture

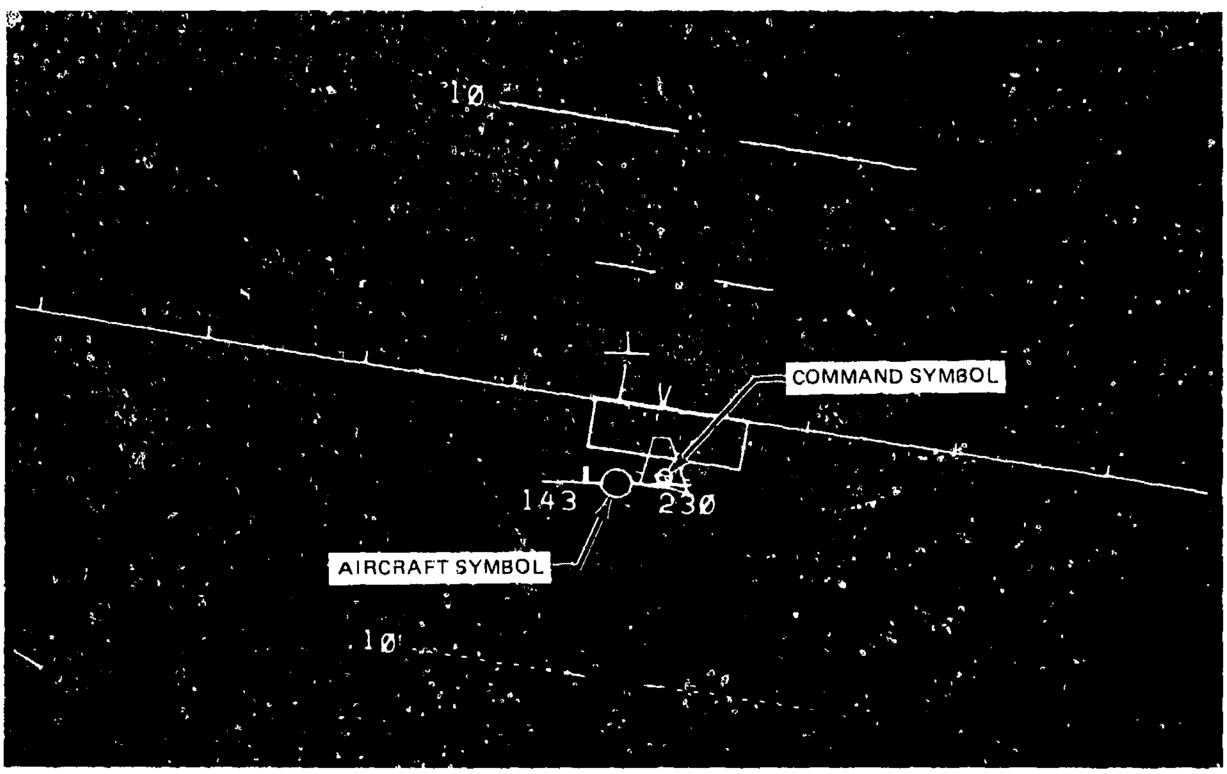


Figure A.2.2-7. HUD Symbology Used for Tests III and IV

APPENDIX B

SUBJECTIVE TEST INSTRUMENTS

- TEST 5 QUESTIONNAIRE
- DEBRIEFING QUESTIONNAIRES

Observer No. ____________

Name: ______________________________ Date: ______________

Phone: ______________________________

Age: ______________________________

Number of years flying: ____________________

Approximate number of flight hours: ____________

In the space below, identify the types of aircraft you have flown. Put a 1 above the aircraft type you have flown most recently, a 2 above the next, and so on.

____	____	____	____	____	____	____	____	____	____
(B-707)	(B-727)	(B-737)	(B-747)	(DC-9)	(DC-10)	(L-1011)	( )	( )	( )
								Other	

Overflow Concept Evaluation

Instructions: Observer No. ________________

After reviewing the candidate concepts, answer the questions below by putting a mark on the scale for each concept.

Please use the comment section to record any thoughts that you have about the concepts, especially in those areas that you consider of great benefit or in which you have serious problems. If you gave a concept a low rating (4 or below), use the comment section to provide suggestions on how the concept(s) might be improved.

Candidate concepts to be evaluated:

(D) Messages drop off bottom

(R) Ability to roll the messages off the top (or bottom)

(S) Subsystem indications for cautions and advisories

Unacceptable — Excellent

1. How good was overall logic of the overflow concept?

(D) I 2 I 4 I 6 I 8 I 10
(R) I 2 I 4 I 6 I 8 I 10
(S) I 2 I 4 I 6 I 8 I 10

2. How well could you evaluate the aircraft status?

(D) I 2 I 4 I 6 I 8 I 10
(R) I 2 I 4 I 6 I 8 I 10
(S) I 2 I 4 I 6 I 8 I 10

3. How good is the level of detail in the messages?

(D) I 2 I 4 I 6 I 8 I 10
(R) I 2 I 4 I 6 I 8 I 10
(S) I 2 I 4 I 6 I 8 I 10

4. How well does concept avoid confusion about priority level, alert position, number of alerts, and overall status?

(D) I 2 I 4 I 6 I 8 I 10
(R) I 2 I 4 I 6 I 8 I 10
(S) I 2 I 4 I 6 I 8 I 10

5. How do you rate the concept in decreasing probability of error?

(D) I 2 I 4 I 6 I 8 I 10
(R) I 2 I 4 I 6 I 8 I 10
(S) I 2 I 4 I 6 I 8 I 10

6. Overall, how easy is the concept to use and understand?

(D)	I	2	I	4	I	6	I	8	I	10
(R)	I	2	I	4	I	6	I	8	I	10
(S)	I	2	I	4	I	6	I	8	I	10

7. How well does concept aid in speed of detecting the most recent alert for the first time?

(D)	I	2	I	4	I	6	I	8	I	10
(R)	I	2	I	4	I	6	I	8	I	10
(S)	I	2	I	4	I	6	I	8	I	10

8. How well can you identify the chronological order of alerts?

(D)	I	2	I	4	I	6	I	8	I	10
(R)	I	2	I	4	I	6	I	8	I	10
(S)	I	2	I	4	I	6	I	8	I	10

9. How well does concept handle the possibility of losing track of an alert?

(D)	I	2	I	4	I	6	I	8	I	10
(R)	I	2	I	4	I	6	I	8	I	10
(S)	I	2	I	4	I	6	I	8	I	10

Comments:

Color Concept Evaluation

Instructions:

Use the letters (A) for amber and (B) for blue to designate where the colors fall on the scale.

Unacceptable Excellent

1. How well can you distinguish the alert priority level?

(A) | 2 | 4 | 6 | 8 | 10
(B) | 2 | 4 | 6 | 8 | 10

2. How well would you be able to evaluate the aircraft status?

(A) | 2 | 4 | 6 | 8 | 10
(B) | 2 | 4 | 6 | 8 | 10

3. How easy is the display to use?

(A) | 2 | 4 | 6 | 8 | 10
(B) | 2 | 4 | 6 | 8 | 10

4. How does the color aid in detecting the most recent alert?

(A) | 2 | 4 | 6 | 8 | 10
(B) | 2 | 4 | 6 | 8 | 10

5. How well does the concept handle the possibility of confusing alert priority levels?

(A) | 2 | 4 | 6 | 8 | 10
(B) | 2 | 4 | 6 | 8 | 10

6. How well does the concept handle the possibility of making errors?

(A) | 2 | 4 | 6 | 8 | 10
(B) | 2 | 4 | 6 | 8 | 10

Comments:

Additional Information

1. Reviewing the three overflow concepts, what features would you add to these concepts to improve system performance?

2. If you would propose a different way to handle the overflow situation, please indicate it here.

3. What do you feel about the usefulness of displaying alerts in chronological order to provide an indication of the sequence in which events occurred? Would you prefer some other format? If so, what and why?

4. Additional comments:

VISUAL DISPLAY
CONCEPT EVALUATION

PART I
OVERFLOW LOGIC CONCEPTS

INSTRUCTIONS: FOR EACH OF THE OVERFLOW LOGIC CONCEPTS DEFINED BELOW, RATE THE THE SYSTEM EFFECTIVENESS BY ASSIGNING A SCALE VALUE FROM 1 TO 5. CIRCLE THE APPROPRIATE NUMBER FOR EACH DISPLAY CHARACTERISTIC. REFER TO PICTORIAL DESCRIPTIONS OF EACH CONCEPT AS NECESSARY.

Candidate concepts to be evaluated:

(D) Messages drop off bottom.

(R) Ability to roll messages off the top (or bottom)

(S) Subsystem indications for cautions and advisories.

Scoring Options:

1 = Excellent -- No changes recommended.
2 = Good -- Minor changes beneficial.
3 = Fair -- Minor changes recommended.
4 = Poor -- Major changes recommended.
5 = Unacceptable -- Major changes necessary.

NOTE: Ratings of 4 and 5 indicate that you anticipate significant operational problems.

1. How good was the overall logic of the overflow concept?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

2. How well would you be able to evaluate the aircraft status?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

3. How good is the level of detail in the messages?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

4. How good is the concept at helping to avoid confusion about priority level?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

5. How good is the concept at helping to avoid confusion about number of alerts?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

6. How good is the concept at helping to avoid confusion about overall status?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

7. How good is the probability of avoiding errors with the concept?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

8. Overall, how easy is the concept to use and understand?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

9. How well does the concept aid in the speed of detecting the most recent alert for the first time?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

10. How well can you identify the chronological order of the alerts?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

11. How well does the concept handle the possibility of losing track of an alert?

(D)	1	2	3	4	5
(R)	1	2	3	4	5
(S)	1	2	3	4	5

Comments:

VISUAL DISPLAY
CONCEPT EVALUATION

PART II
COLOR CODING SCHEMES

INSTRUCTIONS: FOR EACH OF THE COLOR CODING SCHEMES DEFINED BELOW, RATE THE SYSTEM EFFECTIVENESS BY ASSIGNING A SCALE VALUE FROM 1 TO 5. CIRCLE THE APPROPRIATE NUMBER FOR EACH DISPLAY CHARACTERISTIC. REFER TO PICTORIAL DESCRIPTIONS OF EACH CONCEPT AS NECESSARY.

Color coding schemes to be evaluated:

(A) Advisories presented in same color as cautions (amber) and identified by one-space indent.
(B) Advisories differentiated from cautions by blue color code and one-space indent.

Scoring options:

1. = Excellent – No changes recommended.
2 = Good – Minor changes beneficial.
3 = Fair – Minor changes recommended.
4 = Poor – Major changes recommended.
5 = Unacceptable – Major changes necessary.

NOTE: Ratings of 4 and 5 indicate that you anticipate significant operational problems.

1. How well can you distinguish the alert priority level?

(A)	1	2	3	4	5
(B)	1	2	3	4	5

2. How well would you be able to evaluate the aircraft status?

(A)	1	2	3	4	5
(B)	1	2	3	4	5

3. How easy is the display to use?

(A)	1	2	3	4	5
(B)	1	2	3	4	5

4. How well does the concept handle the possibility of confusing alert priority levels?

(A)	1	2	3	4	5
(B)	1	2	3	4	5

5. How well does the concept handle the possibility of making errors?

(A)	1	2	3	4	5
(B)	1	2	3	4	5

Comments:

VISUAL DISPLAY
CONCEPT EVALUATION

PART III
ADDITIONAL INFORMATION

1. Reviewing the three overflow concepts, what features would you add to these concepts to improve the system performance?

____

____

____

____

____

2. If you would propose a different way to handle the overflow situation please indicate it here.

____

____

____

____

____

3. What do you feel about the usefulness of displaying the alerts in chronological order to provide an indication of the sequence in which the events occurred? Would you prefer some other format? If so what and why?

____

____

____

____

____

4. Additional Comments:

____

____

____

____

____

ALERTING TONE EVALUATION QUESTIONNAIRE

NAME ______________________ DATE ____________

ORGANIZATION REPRESENTED ______________________

EXPERIENCE IN TRANSPORT AIRCRAFT (HOURS) ____________ AGE ____________

I. IN THE SPACES BELOW, IDENTIFY THE TYPES OF AIRCRAFT YOU HAVE FLOWN. PUT A "1" ABOVE THE AIRCRAFT TYPE YOU HAVE FLOWN MOST RECENTLY, A "2" BY THE NEXT, ETC.

___	___	___	___	___	___	___	___	___
(DC-9)	(DC-10)	(B707)	(B727)	(B737)	(B747)	(L-1011)	( )	( )
							OTHER	

II. AFTER HEARING EACH OF THE FOLLOWING TONES, BRIEFLY IDENTIFY ANY SPECIFIC MEANING THAT YOU ASSOCIATE WITH IT. THE TONES WILL BE PRESENTED APPROXIMATELY SIX SECONDS APART, SO IT IS IMPORTANT THAT YOUR RESPONSES ARE BRIEF.

TONE NO. 1 ____________ TONE NO. 7 ____________

TONE NO. 2 ____________ TONE NO. 8 ____________

TONE NO. 3 ____________ TONE NO. 9 ____________

TONE NO. 4 ____________ TONE NO. 10 ____________

TONE NO. 5 ____________ TONE NO. 11 ____________

TONE NO. 6 ____________ TONE NO. 12 ____________

III. YOU WILL NOW HEAR A SERIES OF PAIRED TONES, PRESENTED AGAINST A SIMULATED COCKPIT NOISE BACKGROUND. FOR EACH PAIR, PLACE A CHECK (✓) MARK IN THE SPACE REPRESENTING THE TONE THAT YOU FEEL HAS THE <u>BEST</u> ATTENTION-GETTING QUALITY. THE PAIRS WILL BE PRESENTED APPROXIMATELY FOUR SECONDS APART. MAKE YOUR RESPONSES IMMEDIATELY TO AVOID MISSING A TONE OR LOSING YOUR PLACE.

	FIRST	SECOND		FIRST	SECOND		FIRST	SECOND		FIRST	SECOND		FIRST	SECOND
1.	___	___	14.	___	___	27.	___	___	40.	___	___	53.	___	___
2.	___	___	15.	___	___	28.	___	___	41.	___	___	54.	___	___
3.	___	___	16.	___	___	29.	___	___	42.	___	___	55.	___	___
4.	___	___	17.	___	___	30.	___	___	43.	___	___	56.	___	___
5.	___	___	18.	___	___	31.	___	___	44.	___	___	57.	___	___
6.	___	___	19.	___	___	32.	___	___	45.	___	___	58.	___	___
7.	___	___	20.	___	___	33.	___	___	46.	___	___	59.	___	___
8.	___	___	21.	___	___	34.	___	___	47.	___	___	60.	___	___
9.	___	___	22.	___	___	35.	___	___	48.	___	___	61.	___	___
10.	___	___	23.	___	___	36.	___	___	49.	___	___	62.	___	___
11.	___	___	24.	___	___	37.	___	___	50.	___	___	63.	___	___
12.	___	___	25.	___	___	38.	___	___	51.	___	___	64.	___	___
13.	___	___	26.	___	___	39.	___	___	52.	___	___	65.	___	___
												66.	___	___

IV. THE NEXT SERIES OF PAIRED TONES WILL BE RATED FOR THEIR LEVEL OF ANNOYANCE. FOR EACH PAIR OF TONES, PLACE A CHECK (✓) MARK IN THE SPACE REPRESENTING THE TONE THAT YOU FIND TO BE MOST ANNOYING.

	FIRST	SECOND		FIRST	SECOND		FIRST	SECOND		FIRST	SECOND		FIRST	SECOND
1.			14.			27.			40.			53.		
2.			15.			28.			41.			54.		
3.			16.			29.			42.			55.		
4.			17.			30.			43.			56.		
5.			18.			31.			44.			57.		
6.			19.			32.			45.			58.		
7.			20.			33.			46.			59.		
8.			21.			34.			47.			60.		
9.			22.			35.			48.			61.		
10.			23.			36.			49.			62.		
11.			24.			37.			50.			63.		
12.			25.			38.			51.			64.		
13.			26.			39.			52.			65.		
												66.		

V. THE FOLLOWING CRITERIA DEFINE THREE DISTINCT LEVELS OF ALERT URGENCY:

WARNING: EMERGENCY OPERATIONAL OR AIRCRAFT SYSTEMS CONDITIONS WHICH REQUIRE IMMEDIATE CORRECTIVE OR COMPENSATORY ACTION BY THE CREW.

CAUTION: ABNORMAL OPERATIONAL OR AIRCRAFT SYSTEMS CONDITIONS WHICH REQUIRE IMMEDIATE CREW AWARENESS AND SUBSEQUENT CORRECTIVE OR COMPENSATORY CREW ACTION.

ADVISORY: OPERATIONAL OR AIRCRAFT SYSTEMS CONDITIONS WHICH REQUIRE CREW AWARENESS AND MAY REQUIRE CREW ACTION.

FOR EACH OF THE FOLLOWING TONES, DECIDE WHICH LEVEL OF URGENCY IS BEST REPRESENTED BY THAT TONE. CHECK THE APPROPRIATE BOX.

TONE NO.	WARNING	CAUTION	ADVISORY
1			
2			
3			
4			
5			
6			

TONE NO.	WARNING	CAUTION	ADVISORY
7			
8			
9			
10			
11			
12			

VI. IN THE SPACES BELOW, LIST ANY OPERATIONAL OR AIRCRAFT CONDITONS THAT SHOULD BE ANNUNCIATED BY SPECIFIC TONES. FOR PURPOSES OF ANSWERING THIS QUESTION, THE FOLLOWING ASSUMPTIONS SHOULD BE MADE:

A. WARNING, CAUTION, AND ADVISORY LEVEL TONES HAVE BEEN SELECTED

B. AURAL ALERTS FOR SPECIFIC FAILURES WILL BE LIMITED TO A MAXIMUM OF FIVE (5) DISCRETE TONES.

WHAT CONDITIONS (IF ANY) SHOULD HAVE SPECIFIC AURAL ALERTS?	WHAT TYPE OF TONE SHOULD BE USED FOR EACH?
1.	
2.	
3.	
4.	
5.	

Debriefing Questionnaire
Master Alert Evaluation

Instructions:

Observer No. ______________

During test flight, the alerting system was configured using a number of different concepts. Please answer the following questions about alerting concepts. Read instructions carefully for each section so that you are familiar with the code being used for each alternative concept. While you are filling out the questionnaires, the test conductor will be available for questions and the set of photos may be used for easy reference. After each question there is a rating scale (1 to 10) on which you should indicate a rating for each concept.

Please use the comment section to record any thoughts that you have about the concepts, especially in those areas that you consider of great benefit or in which you have serious problems. If you gave a concept a low rating (4 or below), use the comment section to provide suggestions on how the concept(s) might be improved.

Candidate concepts to be evaluated:

(S) Single system in which the same light goes on for both warnings and cautions

(D) Dual system with a separate light for warnings and cautions

(F) Dual system as described above with alert flashing

Unacceptable — Excellent

1. How good was the attention-getting quality of alert?

(S) | 2 | 4 | 6 | 8 | 10
(D) | 2 | 4 | 6 | 8 | 10
(F) | 2 | 4 | 6 | 8 | 10

2. Did master alert provide enough information?

(S) | 2 | 4 | 6 | 8 | 10
(D) | 2 | 4 | 6 | 8 | 10
(F) | 2 | 4 | 6 | 8 | 10

3. Was location of alert appropriate during all flight tasks?

(S) | 2 | 4 | 6 | 8 | 10
(D) | 2 | 4 | 6 | 8 | 10
(F) | 2 | 4 | 6 | 8 | 10

4. How well did concept avoid confusion?

(S) | 2 | 4 | 6 | 8 | 10
(D) | 2 | 4 | 6 | 8 | 10
(F) | 2 | 4 | 6 | 8 | 10

5. How well did concept avoid possibility of pilot error?

(S) | 2 | 4 | 6 | 8 | 10
(D) | 2 | 4 | 6 | 8 | 10
(F) | 2 | 4 | 6 | 8 | 10

6. Did concept avoid distractions that might disrupt performance of flight duties?

(S)	I	2	I	4	I	6	I	8	I	10
(D)	I	2	I	4	I	6	I	8	I	10
(F)	I	2	I	4	I	6	I	8	I	10

Comments:

Alerting Message Formats

Instructions:

For this section use the following code to rate alternative concepts for the message format.

Candidate concepts to be evaluated:

(A) All alerts in separate categories (grouping)

(W) Warnings in separate category, cautions and advisories combined in chronological order

(C) All alerts combined in chronological order

Unacceptable — Excellent

1. How good was the overall logic of the overflow concept?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

2. How well could you evaluate the aircraft status?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

3. How good is the level of detail in the messages?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

4. How well does the concept avoid confusion about priority level, alert position, number of alerts, and overall status?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

5. How do you rate the concept in decreasing probability of error?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

6. Overall, how easy is the concept to use and understand?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

7. How well does the concept aid in speed of detecting the most recent alert for the first time?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

8. How well can you identify the chronological order of alerts?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

9. How well does concept handle the possibility of losing track of an alert?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

Comments:

Additional Information

1. Reviewing the three master alerting concepts, what features would you add to these concepts to improve system performance?

2. What feature would you add to the message format logic to improve system performance?

3. What do you feel about the usefulness of displaying alerts in chronological order to provide an indication of the sequence in which events occurred? Do you prefer some other format? If so, what and why?

4. Additional comments:

System Feature Identification

Instructions:

Considering all the alerting concepts that have been used during your test flight, please identify below five features of the visual alerting system that you liked best:

1.

2.

3.

4.

5.

Five features that you liked least:

1.

2.

3.

4.

5.

Five changes that you would make (if any):

1.

2.

3.

4.

5.

Debriefing Questionnaire
Master Alerts and Flashing Box

Instructions: Observer No. ____________

During test flight, the alerting system was configured using a number of different concepts. Please answer the following questions about alerting concepts. Read instructions carefully for each section so that you are familiar with the code being used for each alternative concept. While you are filling out the questionnaires, the test conductor will be available for questions and the set of photos may be used for easy reference. After each question there is a rating scale (1 to 10) on which you should indicate a rating for each concept.

Please use the comment section to record any thoughts that you have about the concepts, especially in those areas that you consider of great benefit or in which you have serious problems. If you gave a concept a low rating (4 or below), use the comment section to provide suggestions on how the concept(s) might be improved.

Candidate concepts to be evaluated:

(M) Master alert alone without flashing box around message

(B) Flashing box on screen in front of pilot for both warnings and caution without master alert

(C) Flashing box on center screen for both warning and caution without master alert

Unacceptable — Excellent

1. How good was the attention-getting quality of alert?

(M)	ı	2	ı	4	ı	6	ı	8	ı	10
(B)	ı	2	ı	4	ı	6	ı	8	ı	10
(C)	ı	2	ı	4	ı	6	ı	8	ı	10

2. Did alert provide enough information?

(M)	ı	2	ı	4	ı	6	ı	8	ı	10
(B)	ı	2	ı	4	ı	6	ı	8	ı	10
(C)	ı	2	ı	4	ı	6	ı	8	ı	10

3. Was location of alert appropriate during all flight tasks?

(M)	ı	2	ı	4	ı	6	ı	8	ı	10
(B)	ı	2	ı	4	ı	6	ı	8	ı	10
(C)	ı	2	ı	4	ı	6	ı	8	ı	10

4. How well did concept avoid confusion?

(M)	ı	2	ı	4	ı	6	ı	8	ı	10
(B)	ı	2	ı	4	ı	6	ı	8	ı	10
(C)	ı	2	ı	4	ı	6	ı	8	ı	10

5. How well did concept avoid possibility of pilot error?

(M)	ı	2	ı	4	ı	6	ı	8	ı	10
(B)	ı	2	ı	4	ı	6	ı	8	ı	10
(C)	ı	2	ı	4	ı	6	ı	8	ı	10

6. Did concept avoid distractions that might disrupt performance of flight duties?

(M)	ı	2	ı	4	ı	6	ı	8	ı	10
(B)	ı	2	ı	4	ı	6	ı	8	ı	10
(C)	ı	2	ı	4	ı	6	ı	8	ı	10

Comments:

Alerting Message Location

Instructions:

For this section use the following code to rate alternative concepts for the message format.

(A) All alerts on center screen

(W) Warnings in front of pilot, cautions and advisories on center screen

(C) Cautions and warnings in front of pilot and advisories on center screen

Unacceptable — Excellent

1. How good was the overall logic of the location concept?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

2. How well could you evaluate the aircraft status?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

3. How good is the level of detail in the messages?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

4. How well does concept avoid confusion about priority level, alert position, number of alerts, and overall status?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

5. How do you rate the concept in decreasing probability of error?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

6. Overall, how easy is the concept to use and understand?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

7. How well does concept aid in speed of detecting the most recent alert for the first time?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

8. How well can you identify the chronological order of alerts?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

9. How well does concept handle the possibility of losing track of an alert?

(A)	I	2	I	4	I	6	I	8	I	10
(W)	I	2	I	4	I	6	I	8	I	10
(C)	I	2	I	4	I	6	I	8	I	10

Comments:

System Feature Identification

Instructions:

Considering all the alerting concepts that have been used during your test flight, please identify below five features of the visual alerting system that you liked best:

1.

2.

3.

4.

5.

Five features that you liked least:

1.

2.

3.

4.

5.

Five changes that you would make (if any):

1.

2.

3.

4.

5.

Additional Information

1. Reviewing the three master alerting concepts, what features would you add to these concepts to improve system performance?

2. What feature would you add to the message location logic to improve system performance?

3. What do you feel about the usefulness of displaying alerts in chronological order to provide an indication of the sequence in which events occurred? Do you prefer some other format? If so, what and why?

4. Additional comments:

TEST III
DEBRIEFING CHECKLIST

SUBJECT ______________________________ DATE __________

	Word-Phrase	About Equal	Complete Sentence
1. Which voice alert format was easiest to identify when presented simultaneously with an ATC message?	☐	☐	☐
2. Which voice alert format caused most difficulty in identifying concurrent ATC messages?	☐	☐	☐
3. Which voice alert format would you prefer for a cockpit alerting system?	☐	☐	☐

	Voice Only	About Equal	Tone-Voice
4. Which alerting mode was most effective in getting your attention?	☐	☐	☐
5. Which alerting mode would you prefer for a cockpit warning system?	☐	☐	☐

	Too Low	About Right	Too High
6. Evaluate the alert loudness (with respect to aircraft noise).	☐	☐	☐

	Unacceptable	Satisfactory	Excellent
7. Evaluate the alerting tones selected for this study.			
Warning Horn	☐	☐	☐
Caution C-chord	☐	☐	☐

	Yes	No
8. Do you anticipate significant operational problems with confusion of voice alerts and crew/ATC communications?	☐	☐

9. What provisions should be made for cancellation/attenuation of auditory alerts?

	Check one or more
(a) Manual Cancellation (central switch)	☐
(b) Manual Cancellation (dedicated switch)	☐
(c) Automatic cancel after fixed number of repetitions	☐
(d) Automatic volume reduction after fixed number of repetitions	☐
(e) Cancel only when problem corrected	☐
(f) Depends on the nature of the alert	☐

Comments:

____

____

____

____

TEST IV
DEBRIEFING CHECKLIST

SUBJECT ______________________________ DATE __________

	Tone-Visual	Tone-Voice	Voice Only	Tone-Voice-Visual
1. Which alerting mode was most effective in getting your attention?	☐	☐	☐	☐
2. Which alerting mode would you prefer for a cockpit warning system?	☐	☐	☐	☐

	Too Low	About Right	Too High
3. Evaluate the alert loudness (with respect to aircraft noise).	☐	☐	☐

	Unacceptable	Satisfactory	Excellent
4. Evaluate the alerting tones selected for this study.			
Warning Horn	☐	☐	☐
Caution C-chord	☐	☐	☐

	Yes	No
5. Do you anticipate significant operational problems with confusion of voice alerts and crew/ATC communications?	☐	☐

6. What provisions should be made for cancellation/attenuation of auditory alerts?

	Check one or more
(a) Manual Cancellation (central switch)	☐
(b) Manual Cancellation (dedicated switch)	☐
(c) Automatic cancel after fixed number of repetitions	☐
(d) Automatic volume reduction after fixed number of repetitions	☐
(e) Cancel only when problem corrected	☐
(f) Depends on the nature of the alert	☐

Comments:

____

____

____

____

____

____

____

APPENDIX C

TESTS 3 AND 4

ATC TRANSMISSIONS AND ALERT MESSAGE CONSTRUCTION

ATC MESSAGE FORMATS

	MESSAGE	PEAK SOUND LEVEL (dBA)* MALE	PEAK SOUND LEVEL (dBA)* FEMALE	DURATION (SECONDS) MALE	DURATION (SECONDS) FEMALE
(CLEARANCE)	DACO 891 CLEARED TO LAND RUNWAY TWO-FIVE LEFT WIND TWO-SEVEN ZERO AT ONE-TWO	85	85	3.5	3.8
(1)	DACO 891 TRAFFIC TWO O'CLOCK TWO MILES SOUTHBOUND, ALTITUDE UNKNOWN	85	85	3.7	3.8
(2)	DACO 891 BRAKING ACTION REPORTED FAIR BY A DC-9	83	86	3.1	3.1
(3)	DACO 891 WINDSHEAR REPORTED AT 400 FEET 1/2 MILE FROM RUNWAY	85	86	3.7	4.2
(4)	DACO 891 CHANGE TO RUNWAY TWO-FIVE RIGHT, CLEARED TO LAND WIND TWO-SEVEN-ZERO AT ONE-TWO	84	87	3.9	4.4
(5)	DACO 891 ALTITUDE ALERT, CHECK YOUR ALTITUDE IMMEDIATELY	86	85	3.7	3.6
(6)	DACO 891 TRAFFIC TO YOUR RIGHT IS FOR RUNWAY TWO-FIVE RIGHT	86	84	3.0	3.0
(7)	DACO 891 SLOW TO YOUR FINAL APPROACH SPEED, OVERTAKING TRAFFIC AHEAD	85	86	3.6	3.4
(8)	DACO 891 GO AROUND, FLY HEADING TWO-FIVE-ZERO, CLIMB AND MAINTAIN TWO THOUSAND	85	87	4.5	4.5
	MEAN	84.90	85.70	3.63	3.75
	STANDARD DEVIATION	0.93	1.00	0.44	0.54
	RANGE MAXIMUM	86.00	87.00	4.50	4.50
	RANGE MINIMUM	83.00	84.00	3.00	3.00

*NOTE: SOUND LEVEL MEASUREMENTS RECORDED IN THE COCKPIT AT THE PILOT'S EAR POSITION.

VOICE ALERT MESSAGE FORMATS

	WORD-PHRASE FORMAT	PEAK SOUND LEVEL (dBA)*	DURATION (SECONDS)	SENTENCE FORMAT	PEAK SOUND LEVEL (dBA)*	DURATION (SECONDS)
WARNING MESSAGES	LANDING GEAR	85		THE LANDING GEAR IS NOT DOWN	86	
	EMERGENCY BUS	87		THE EMERGENCY BUS HAS FAILED	86	
	FIRE ENGINE 1	85		YOU HAVE FIRE IN ENGINE 1	85	
	FIRE CARGO	86		YOU HAVE FIRE IN CARGO COMPARTMENT	85	
CAUTION MESSAGES	AUTO SPOILERS	8[illegible]		THE AUTO SPOILERS ARE NOT ARMED	88	
	HYDRAULIC FAILURE	84		YOU HAVE HYDRAULIC SYSTEM FAILURE	85	
	FUEL PRESSURE LOW	86		THE FUEL PRESSURE IS LOW	86	
	ANTI-SKID	82		THE ANTI-SKID IS NOT ARMED	83	
	MEAN	85.00		MEAN	85.50	
	STANDARD DEVIATION	1.51		STANDARD DEVIATION	1.41	
	RANGE MAXIMUM	87.00		RANGE MAXIMUM	88.00	
	RANGE MINIMUM	82.00		RANGE MINIMUM	83.00	

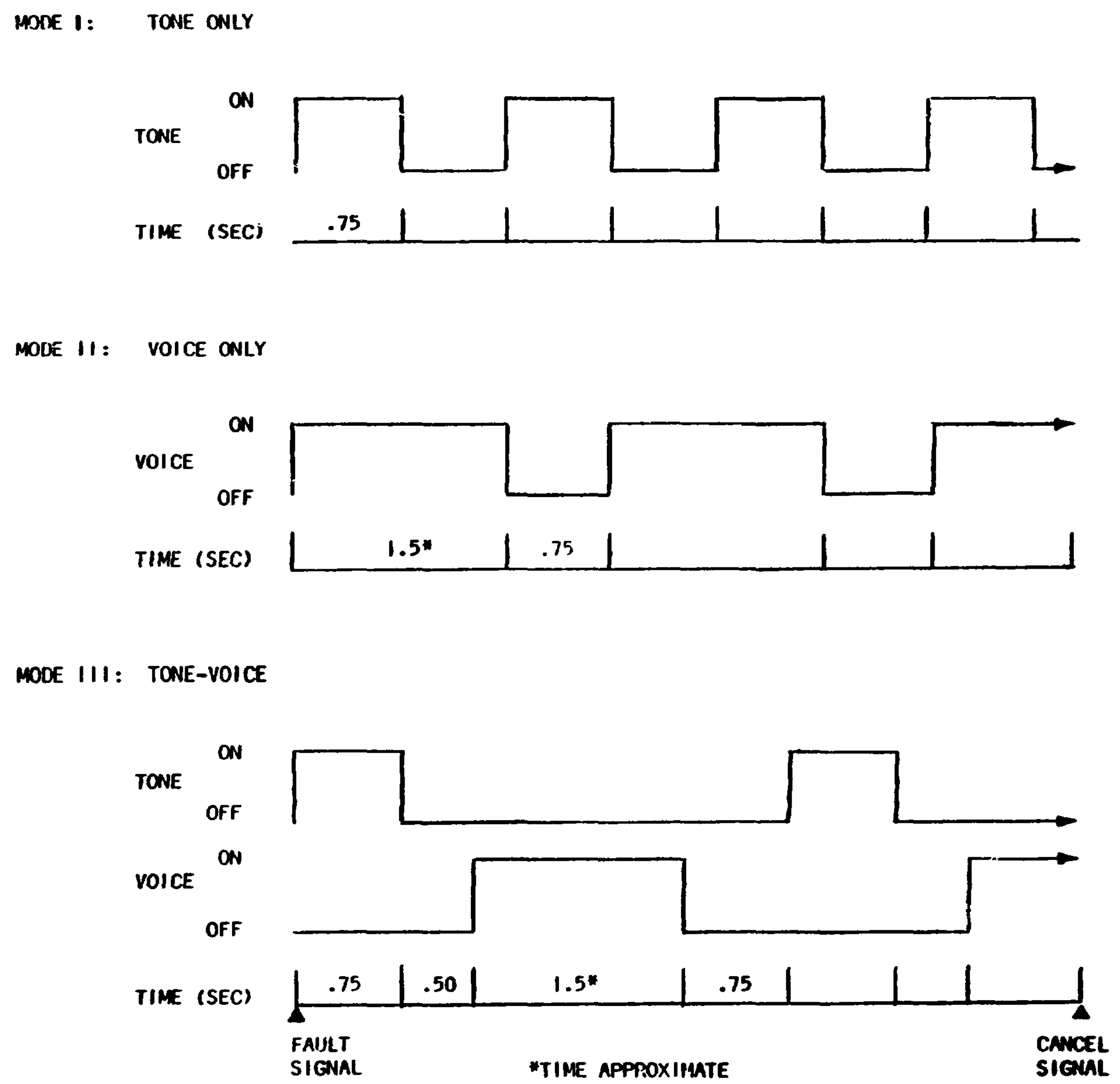
*NOTE: SOUND LEVEL MEASUREMENTS RECORDED IN THE COCKPIT AT THE PILOT'S HEAD POSITION

ALERTING TONE CHARACTERISTICS

	PEAK SOUND LEVEL (dBA)*	DURATION (SECONDS)
WARNING HORN	84	0.75
CAUTION C-CHORD	83	0.75

*NOTE: SOUND LEVEL MEASUREMENTS RECORDED IN THE COCKPIT AT THE PILOT'S HEAD POSITION

AURAL ALERT SEQUENCING



APPENDIX D

TEST INSTRUCTIONS

TEST SESSION INSTRUCTIONS - TESTS 1 AND 2

During the test sessions you will be required to perform a number of tasks. These will include: following the flight director and reference speed director using the wheel, column and throttles; responding to ATC queries about certain flight parameters; identifying targets on the three screens in front of the cab; responding to alerts as they occur during the flight; and completing questionnaires about the alerting system. Data will be gathered on all these tasks to determine their effect on each other.

The simulation that you will be flying is not meant to realistically represent any existing aircraft or any specific flight phase. It is designed to provide a workload and work pattern that is similar to those encountered in the flight situation. The outside visual scene is not realistic in terms of actual flight but it does make you go head up and focus at infinity.

The alerts that will be appearing may not be appropriate at that point in time and the response to the alert will be different than what you are familiar doing. However, it is important to keep in mind that we are looking at the alerting system itself rather than the specific alerts. Furthermore, each individual may have slightly different procedural responses to some of the alerts in actual flight situations and to overcome these differences we have standardized the responses to the alerts. The ATC queries will also be more frequent than would normally occur, however they will aid in simulating some of the mechanical and scan tasks that you perform during your normal flight procedures. Let me stress again that the total simulation is not meant to be an exact or even a close representation of a full flight task but rather it will provide tasks that require manual, visual, auditory and mental work similar to those performed during flight.

Let me explain each of your tasks individually and then we will give you a practice flight to become familiar with the system as a whole. If you have any questions at any time feel free to ask them.

FLIGHT TASK - FLIGHT DIRECTOR AND SPEED DIRECTOR

When you are flying the simulator you will have the wheel/column and the throttle active. Your task will be to follow the flight director in both pitch and roll. The flight path has no navigational base but rather is a computer generated random path in the sky. As you are following the path you will also be required to adjust your airspeed to hold 110 knots on the airspeed indicator. Your flight performance will be a measure of how close you were able to follow the director. The simulated airplane is a low speed plane it must be flown below 145 knots or it becomes unstable. If you exceed this speed during the flight take corrective action.

ATC RESPONSES

Periodically throughout your flight you will be requested by ATC to report on either altitude, or heading. When these calls occur you will be required not only to report verbally but also to input your response on this (show keyboard) panel by pushing the appropriate whole numbers and the enter (E) key. If you should make a mistake push the clear (C) button and begin again. Your score for this task will be a combination of your entry speed and accuracy of response.

OUTSIDE VISUAL

Also periodically throughout your flight ATC will inform you that you have simulated traffic in the area. When you hear this report slides containing broken circles will come up on all the three screens (show example). Within each screen all the circles will be oriented in the same direction. On one of the screens will be a circle which is oriented in a different direction. Your task will be to find the odd circle and report its position. You will do this by inputting again into your 10-key keyboard a two digit number. The first digit will correspond to the screen number and they are numbered from left to right (1, 2, 3). The second digit will correspond to the quadrant on the screen numbered from the upper left quadrant as shown on the diagram in front of you. If your target seems to be on the line between two quadrants try to estimate in which quadrant the greater portion of it lies.

ALERT RESPONSES

During the flight different types of alerts will be activated. These will be three basic priority levels:

Warnings - which are defined as conditions which require immediate action.

Cautions - which are defined as conditions which require immediate awareness.

Advisories - which are defined as conditions which require awareness.

These alerts will appear in different formats depending on the test conditions and you will be informed before each flight what configuration the alerting system will assume. Your task for all of the systems will be the same. When you detect that an alert has occurred you depress the trigger under your left hand and announce that you have detected a "warning", "caution" or "advisory". In order to avoid procedural problems that might occur because of your previous aircraft experience we have devised a very simple response to the alerts. Once you have identified the alert and recognized the system in which you have a problem you will respond when you feel it is appropriate to do so by depressing the appropriate system switch on this overhead panel (demonstrate). If your response is correct the master alert will extinguish and the message will be erased. If you feel more comfortable cancelling the master alert before you respond you may do so but it is not necessary.

INSTRUCTIONS FOR OVERFLOW LOGIC STUDY

PICTURE 1

Under certain circumstances it is possible for the display to fill up with alerts. This picture shows this situation using a three (two) color scheme to differentiate between the levels of priority red = warning, amber = caution, blue/(amber) = advisory. The advisories are colored blue (colored amber same as the cautions) and indented one space to distinguish them from the cautions. If another alert should appear when the screen is full there are a number of ways the system could handle the overflow condition, each of which has some advantages and disadvantages. The overflow concepts being considered, however, are those which affect only the caution and advisories. The warning messages must be available to the crew at all times. A second general ground rule for the photographs is that the baseline system will have a chronological format for the cautions and advisories which may not be the case for the final system. This means that the most recent message comes in at the top of the cautions and advisories (demonstrate with picture).

Concept 1 (Picture 1, 2, & 3) Straight Overflow

PICTURE 2

The first concept to be considered is the straight overflow. As can be seen here "ANTI-ICE" has appeared on the display as the most recent alert. This caused "L FUEL FWD PUMP" to drop off the bottom of the display. Therefore, the oldest messages, no matter what level of priority (caution or advisory), will drop off.

PICTURE 3

When a fault has been corrected, as we see in this picture "L PACK TRIP" is no longer displayed, all the messages below it move up, returning "L FUEL FWD PUMP" to the display. There will be no blank spaces between messages and no limit to the number of alerts that are held in the buffer off the bottom of

the display. The alerts will re-enter the bottom of the display in the reverse order of their leaving, i.e., the last alert to leave will be the first to re-enter.

Concept 2 (Picture 1, 4, & 5) Roll Mechanism

PICTURE 4

Another way to handle the overflow, and still provide access to all the alerts, is to make some means available to roll, or scroll, the caution and advisory alerts either up or down. This picture shows that the "ANTI-ICE" alert has been rolled off the top of the display to get back the "L FUEL FWD PUMP" message which had been displayed off the bottom. This process would be repeated for other messages that have gone off the bottom.

PICTURE 5

The reverse process is used to roll the messages down; as you see "ANTI-ICE" has reappeared at the top and "L FUEL FWD PUMP" has gone off the bottom.

Concept 3 (Picture 1 & 6) Subsystem Reversion

Finally, since overflow is generally caused by multiple failures in a single system (Note: 3 electrical failures), the display could revert to a system designation -Picture 6-. This would combine all of one system's problems (cautions and advisories) into a single indication. the bottom five lines are reserved for the 10 systems messages. These messages will always appear in the same location. When the number of individual alerts falls below the maximum number of lines in the display, the system will automatically revert back to the normal, full message presentation.

APPENDIX E

PILOTS COMMENTS

COMMENTS

VISUAL DISPLAY CHARACTERISTICS
SUMMARY OF COMMENTS*
(n = 28)

I. OVERFLOW LOGIC EVALUATION	FREQUENCY
A. The Roll concept has a distinct advantage over the Drop-off in that the latter has no provision for system review.	12
B. When an overflow condition occurs there should be some reminder that one or more items are no longer displayed.	9
C. Illumination of the directional keys should be used to indicate an overflow condition with the Roll concept. Keys should not be illuminated for reading purposes.	3
D. An overflow condition should be indicated by a symbol on the primary display. One concept would employ an arrow (↓ ↑) similar to that used on the present RNAV display.	12
E. With the Roll concept, the display of a new item presents a problem when the display is rolled up. If the new item comes in at the top of the display, the chronological order is interrupted. If the new item is assimilated into the chronology scheme, the pilot may not be aware that a new failure has occurred.	3
F. When a new advisory level item is introduced and a display overflow condition exists the advisory message may cause a high priority item to drop-off the bottom of the display.	9

* Comments are paraphrased from the actual pilot comments.

		FREQUENCY
G.	If the Subsystem concept is used, there must be some method provided to call up detailed information for each of the system summary messages.	10

II. COLOR CODING SCHEMES

A.	The use of a single color to represent both cautions and advisories is unacceptable.	13
B.	The identification of advisories by using a one-space indent provides inadequate distinction between caution and advisory levels.	4
C.	White letters should be used to represent advisory level messages. Blue is difficult to read.	1

III. INTERACTIVE FUNCTIONS

A.	Some mechanism for selective cancel/recall of an alert message is necessary.	11
B.	The Roll keys should function in discrete steps. moving one message line for each press of the key.	3
C.	The up/down roll functions should be integrated into a single center-loaded rocker switch.	1
D.	A Page function would be preferred over the roll mechanism for providing access to deferred items.	12
E.	A manual brightness control should be provided.	1

IV. VISUAL DISPLAY CHARACTERISTICS (GENERAL)

		FREQUENCY
A.	There should be some way to determine readily whether an alert message has been attended to.	3
B.	Additional space between lines of text would be desirable.	1
C.	A positive indication of display status change (e.g., new alert) is needed.	10
D.	Any new item on the display should be identified by flashing the message for a fixed duration or number of repetitions. This would be essential if a priority scheme is used.	1
E.	The most recent fault should be identified with an asterisk or other symbol in the margin.	9
F.	The distinction between priority levels would be enhanced by leaving a blank space between warnings and cautions.	1
G.	Information should be grouped <u>and</u> color coded according to priority Within each category items should be displayed in chronological order.	12
H.	The chronological format provides useful information regarding the sequence of events and failure trends. This format assists in establishing causal relationships.	3
I.	Warnings should be included in the chronology scheme.	1
J.	Flashing master warning and caution lights would be preferred over steady-state lights.	1

AUDITORY SYSTEM CHARACTERISTICS
SUMMARY OF COMMENTS
(n = 16)

I. ALERTING TONE CHARACTERISTICS	FREQUENCY
A. The distinction between the two-tones used to represent warning and caution in the simulator tests was inadequate.	7
B. The meanings assigned to the two alerting tones used in the simulator tests should be reversed. The caution C-Chord sounds more urgent than the Warning Horn.	6
C. The Warning horn had an annoying quality.	1
D. High pitched sounds tend to be associated with greater urgency.	4
E. The alerting tone tended to direct my attention to the alert message and I missed the ATC message.	1
F. When used in combination with a voice alert message, the alerting tone should be presented only at the alert onset. Repetitions of the alert should include the voice message only.	1
G. Alerting tones could be coded for priority by number of repetitions of the tone.	1
II. STEREOTYPED ALERTS	
A. Dedicated alerting tones are needed to identify certain immediate action items such as fire and overspeed.	2

B. Tones which have generally accepted meanings such as the fire bell should be retained. 2

C. The conventional system of using discrete tones for each alert condition is preferred. The total number of tones should be reduced by restricting aural alerts to a limited set or critical failures. 1

III. CENTRAL AURAL WARNING SYSTEM CONTROL OPTIONS

A. All alerting tones should be manually cancellable. Aural alerts constitute a potential distraction after they have been annunciated and correctly identified by the crew. 3

B. Selected auditory alerts should stay on until the problem is corrected. Some alerting tones such as stall, overspeed and GPWS provide feedback when the aircraft is within acceptable limits. 2

C. Manual cancellation of alerting tones could be performed most effectively by using the master warning or caution. 6

D. The present system of dedicated push buttons for aural alert cancellation can lead to delays and confusion due to the time required to search for thhe appropriate button. 2

E. The conventional system of dedicated switches for manual cancellation is desirable since the pilot must acknowledge the specific problem. This makes it unlikely that the pilot will misunderstand the meaning of the aural alert. 1

IV. VOICE MESSAGE FORMATS

A. The voice message should state the specific nature and location of the problem rather than commands or instructions for corrective action. 2

B. The complete sentence message format was easy to hear but completely blocked out the ATC communication when both came on at the same time. 1

V. OPERATIONAL PROBLEMS WITH VOICE WARNINGS

A. Extensive use of voice alerts in the cockpit will cause problems with crew/ATC communications. It is difficult to hear and understand more than one voice message at a time. 2

B. Voice alerts should be used selectively to minimize potential for confusion with other communications. Overuse of voice alerts should definitely be avoided. 4

C. The use of voice alerts is a significant improvement over conventional systems that employ individual alerting tones. There are too many tones to remember. 5

D. The use of voice warnings may interfere with communications in high density arrival and departure areas. 1

E. Voice warnings should be inhibited below about 1000 to 1500 feet. 1

F. Confusion of a voice alert and an ATC communication is a relatively low probability event. 3

SAMPLE OF SAME PILOT COMMENTS VERBATIM

The blinking light is very attention-getting, but is more likely to provoke a "knee jerk" reaction (perhaps the wrong reaction). The single light and dual light would probably be noticed in time for non-perilous items.

The single light for either caution or warning is confusing. Blue advisory lights should not be part of this system (i.e., "engine anti-ice valves open" could be illuminated for entire flight) but if they are: a separate blue master alert would be desirable.

Alerts should be in the pilots' line of sight.

The single light (for both warning and caution) was confusing. Two flashing lights in the same location is not desirable. It must be determined which flashed - the red or yellow.

I believe a "quiet, dark cockpit" is the most desirable. Night operations can be jeopardized by flashing lights or confusing aural alerts, particularly if night vision and the primary task of flying the airplane are interrupted. i.e., if a person was paged in a movie theater, stopping the film and turning the lights up bright would be an extreme interruption.

An accurate and timely reaction is better than a fast one. The primary task of flying the airplane should not be grossly interrupted. My philosophy:

1. Fly the airplane first.

2. Red - the airplane is in peril; be accurate, act quickly, but smooth

3. Amber - get this corrected within about 5 minutes.

4. Blue - normal instrument scan will take care of this. Plan appropriately.

I like the flashing box. Our experience has been that about 10 second flashing is enough. The box should then be retained to indicate a new item.

I also believe that the Audio "Attenson" has a definite role for warnings and at least certain cautions.

I tend to favor seeing Warnings, Cautions and Advisories all in the same location, i.e., to reduce the number of scans.

APPENDIX F

DETAC START UP PROCEDURE

APPENDIX F

DETAC START-UP PROCEDURE

1. INITIALIZE VIDEOTAPE EQUIPMENT

A. Insert yellow plug into wall socket (near door).

B. Remove lens cap from table-mounted camera.

C. Turn power on (set record level to mark).

D. Insert videotape.

E Rewind.

F System is now ready to record. Continue with step G when ready to begin taping.

G. Push "RECORD" button down.

Simultaneously - Push "RECORD" and "FWD" buttons down simultaneously

H. Push "FWD" button down.

I. Turn on light bar.

2. TURN AMPLIFIERS ON

A. Turn noise amplifier on mode setting to "TAPE".

B. Set volume to level indicated by pointer.

C. Set volume on both tape units and balance control to appropriate settings.

D. Set background track to "STEREO" mode.

E. Turn cockpit amplifier on.

F. Set volume to level indicated by pointer. Set sound levels.
CAWS message 3: Landing Gear (1144) = 85 dBA
ATC message 5: Alt. Alert = 85 dBA

3. INITIALIZE MICROPROCESSOR

A. Turn both cockpit switches to "ON" position. (These are located next to cockpit amplifier).

B. At microprocessor control board, located in front of and to the right of the cockpit,

(1) Actuate(?) reset switch up.
(2) Actuate(?) start switch up.
(3) Actuate(?) same reset switch up again.
(4) Actuate(?) power switch up (on board below the nose of the airplane--outside cockpit).

4. INITIALIZE ADVENT PROJECTION SYSTEM

A. Set channel selector to Channel 9.

B. Turn volume control to "ON" position.

5. LOAD TAPE FOR PERFORMANCE DATA OUTPUT

(Cassette deck is next to printer near host computer)

A. Load cassette into tape unit.

B. Press "POWER" button (on left side).

C. Press "MOTION" button (on right side).

6. CYCLE THROUGH AURAL WARNINGS FOR SUBJECT

A. Turn on CAWS box (switch is on left side of box).

B. At teletype machine (in main computer room).

(1) Enter: ;A B (then press "RETURN").

(2) Enter: ;R S , CHECK, DK (then press "RETURN").

(3) At the sound of the bell, system is ready for appropriate signal codes.

7. INITIALIZE CENTRAL COMPUTER/VECTOR GENERAL

A. (1) Set left cockpit toggle switch to center position.

(2) Set center cockpit toggle switch to forward position.

(3) Set right cockpit toggle switch to forward position.

B. (1) Actuate 2 lower left switches to "ON" position.

(2) Actuate upper right power switch to "ON" position.

(3) Push program bus and preview bus (No. 2 only) on upper lefthand section.

C. Initialize program.

S: System Command	U: User Response
KB: Keyboard	LP: Light Pen

(1) U: Turn Vector General to "ON" position (switch located on side of unit.

(2) U: Press "INTERRUPT" key (side keyboard).

(3) S: Enter program name.

(4) U: Enter "WARN" (KB), then hit null key.

(5) S: Lun.

(6) U: Enter "DJ" (KB), (hit null key).

(7) U: Hit "START" (LP).

(8) S: Enter subject's name.

(9) U: Enter subject's name (KB), (hit null key).

(10) S: Are you through?

(11) U: Hit "NO" (LP).

(12) U: Hit select stored case (LP).

(13) U: Hit case that activates desired turbulence level for first practice run (LP).

(14) U: Hit finished (LP).

(15) S: Rewinding tape, reading disk, initializing.

(16) U: When HUD symbology comes onto screen:

NOTE: Step (a) is for Test IV only. For Test III. move on to Step (b).

(a) Set up visual messages (TEST IV ONLY).

(1) Turn on cameras (TV).
(2) Turn on visual messages by pressing speed bug (caution) and heading (warning) buttons in cockpit.
(3) Focus cameras accordingly.
(4) Cycle through visual messages for subject. (Cycle through warning message by turning "CONTROL LAW" switch. Cycle through caution messages by turning "RAW DATA" switch).

(b) Set up camera to projected HUD symbology onto Advent screen.

(1) Remove lens cap.
(2) Turn camera to "ON" position.
(3) Line camera up so that:
    (a) Bottom of throttle indicator is just off screen at top.
    (b) Horizon line covers screen from left to right.
(4) Focus camera (front section of lens).
(5) Set intensity (rear section of lens).

D. Initialize first practice run.

(1) Trim throttles.

(2) Set left cockpit toggle to forward position.

(3) Press graphics above HUD symbology with light pen.

E. When run is completed.

(1) If another run using menu is desired, press "DETECT HERE TO CONTINUE" with light pen.

(2) If runs in cell mode are desired.

(a) Push right toggle in cockpit to center position (to enter cell mode).

(b) Push left toggle to center position (to collect data).

(c) Press "DETECT HERE TO CONTINUE" with light pen.

F. (1) S: Are you through?

(2) U: No (LP).

(3) S: Enter cell number.

(4) U: Enter cell number (KB) (Press null key).

(5) U: Enter T_M (KB) (Press null key).

(6) U: Enter alert message (KB). Before pressing null key, be sure that ATC tapes are inserted and rewound.

(7) When HUD symbology comes up:

(a) Arm tapes.
(b) Initialize run with light pen.

8. TEST TRIALS

A. Start video/audio recording.

B. Shift time generator to zero - note time.

C. Left toggle to center position (collect data).

9. VIDEO/AUDIO PLAYBACK

A. Noise amplifier to "TUNER" mode.

B. Volume to maximum output.

C. Video cassette unit to <u>FWD</u>.